

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 21:34:44 ; Search time 36 Seconds

(without alignments)
151.758 Million cell updates/sec

Title: US-09-828-456-2

Perfect score: 207

Sequence: 1 MNRLIELTGMIVLVSVILL.....VASHIDNYQPEQASVQHK 41

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	207	100.0	41	23	AA52607
2	60.5	29.2	231	22	ABP38752
3	58.5	28.3	203	22	AA682688
4	58.5	28.3	203	23	ABP39600
5	58	28.0	291	21	AA609408
6	58	28.0	344	21	AA609407
7	58	28.0	361	21	AA609406
8	57	27.5	327	22	AAU53139
9	55	26.6	320	23	AB54920
10	54.5	26.3	900	22	AA692354

11	54	26.1	1048	22	ABG20939
12	53.5	25.8	206	21	AA611050
13	53.5	25.8	206	22	AAU34255
14	53.5	25.8	206	22	AAU37083
15	52	25.1	455	22	AAU64449
16	52	25.1	459	22	AB67668
17	51.5	24.9	294	22	AB67023
18	51	24.6	86	20	AA134616
19	51	24.6	404	23	AB695184
20	51	24.6	541	23	AB649247
21	51	24.6	567	21	AA624967
22	51	24.6	799	21	AA624966
23	51	24.6	817	21	AA624965
24	51	24.6	1360	22	AB61753
25	50.5	24.4	355	22	AB67739
26	50.5	24.4	491	23	AB638665
27	50	24.2	232	21	AA638330
28	50	24.2	248	22	AAU30985
29	50	24.2	265	11	AA606909
30	50	24.2	265	18	AA614464
31	50	24.2	578	22	AB67874
32	49.5	23.9	97	22	AAU49588
33	49.5	23.9	240	23	AB692584
34	49.5	23.9	551	23	AB629598
35	49.5	23.9	909	22	AAU03554
36	49.5	23.9	1024	23	AB607504
37	49	23.7	103	21	AA605689
38	49	23.7	137	21	AA633211
39	49	23.7	146	21	AA605688
40	49	23.7	146	23	AB691722
41	49	23.7	154	22	AB662506
42	49	23.7	162	21	AA633210
43	49	23.7	176	21	AA603687
44	49	23.7	258	21	AA633209
45	49	23.7	390	11	AA605475

ALIGNMENTS

RESULT 1	
AA52607	standard; Protein; 41 AA.
AA52607	
08-FEB-2002 (first entry)	
Escherichia coli beta-lactam resistance (BLR) protein.	
Beta-lactam resistance; blr gene; Beta Lactam-358; BLR; transmembrane;	
efflux pump; periplasm; peptidoglycan synthesis inhibition;	
drug discovery; screening; antibiotic; antibacterial; vaccine;	
gene therapy.	
Escherichia coli.	
Key	Location/Qualifiers
Domain	6..26 /label= Transmembrane_domain
WO200177176-A2.	
18-OCT-2001.	
06-APR-2001; 2001MO-US11363.	
06-APR-2000; 2000US-195505P.	
14-JUL-2000; 2000US-218380P.	
(TUFT) TUFTS COLLEGE.	
Levy SB, McMurtry LM;	

Novel human diago
Staphylococcus aur
Staphylococcus aur
Staphylococcus aur
Propionibacterium
Drosophila melanog
Drosophila melanog
Chlamydia pneumonia
Human 158P14 sp1
Listeria monocytog
Plant SDF encoded
Plant SDF encoded
Drosophila melanog
Drosophila melanog
Staphylococcus epi
Human secreted pro
Novel human secret
Intra-acrosomal sp
Acrosomal sperm an
Drosophila melanog
Propionibacterium
Herbicidally activ
Streptococcus poly
Human protein kin
Human GTP-binding
Arabidopsis thalia
Zea mays protein f
Herbicidally activ
Drosophila melanog
Zea mays protein f
Arabidopsis thalia
Zea mays protein f
Chloramphenicol r

XX WPI: 2002-010898/01.
DR N-PSDB; ABA02070.
XX
PT New membrane protein, designated Beta Lactam-358 polypeptides, that
PT affect susceptibility to antibiotics which affect peptidoglycan
PT synthesis in microbes, useful for identifying modulators for treating
PT infections
XX
PS Claim 11; Fig 1A; 104pp; English.
XX
CC This sequence represents a beta-lactam resistance (BLR) protein from
CC *Escherichia coli*. The BLR protein is a 41 amino acid transmembrane
CC protein (also designated Beta Lactam-358 in the specification) with a
CC molecular weight of 4556 daltons which confers resistance to beta-lactam
CC antibiotics. The blr gene which encodes it was initially identified in a
CC 602 bp hypothetical intergenic region of the genome of an antibiotic-
CC susceptible strain of *E. coli*, RM583. In the RM583 strain, the blr gene
CC is disrupted by insertion of the transposon *mln1-TnpHoA*. A clone bearing
CC an ORF (open reading frame)-containing 358 bp segment of the "intergenic"
CC region was found to restore beta-lactam resistance in RM583. Certain
CC amber mutations introduced into the 358 bp clone were found to abolish
CC beta-lactam resistance and resistance could be restored using an
CC amber suppressor, indicating that a protein encoded by the clone was
CC responsible for mediating beta-lactam resistance. It is suggested that
CC the BLR protein may act by increasing a beta-lactamase activity in a
CC cell, possibly being a component of an uncharacterised membrane-bound
CC efflux pump capable of expelling beta-lactams from the periplasm. The
CC BLR protein may be used in methods to identify compounds that modulate
CC antibiotic resistance in a bacterium. The BLR protein, gene, BLR protein
CC homologues, BLR agonists or antagonists, and anti-BLR antibodies may be
CC used to treat bacterial infection, particularly by an organism resistant
CC to antibiotics that affect peptidoglycan synthesis. They may also be
CC used in screening assays, in vaccines, and in diagnostic assays. BLR
CC nucleotides are useful in gene therapy applications to modulate BLR
CC expression, and to detect BLR expression in a biological sample.
SQ Sequence 41 AA;
Query Match 100.0%; Score 207; DB 23; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNRLIELTGWIVLVSVILGLVASHIDNYQPEQSASVQHK 41
DB 1 MNRLIELTGWIVLVSVILGLVASHIDNYQPEQSASVQHK 41
RESULT 2
ABP38752
ID ABP38752 standard; Protein; 231 AA.
XX
AC ABP38752;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3597.
XX
KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX antibacterial; gene therapy.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-0134001.
XX
PR 14-AUG-1997; 97US-055779P.
XX 08-NOV-1997; 97US-064964P.
XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm 1A, Bush D;
PI
XX
DR WPI: 2002-381255/41.
DR N-PSDB; ABN91297.
XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS Disclosure; SEQ ID 3597; 267pp; English.
XX
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
SQ Sequence 231 AA;
Query Match 29.2%; Score 60.5; DB 23; Length 231;
Best Local Similarity 35.0%; Pred. No. 2.5;
Matches 14; Conservative 12; Mismatches 7; Indels 7; Gaps 2;
QY 9 GMIYLVSVILGLVAVV-----SHID--NYQPEQSASVQHK 41
DB 8 GIIIFIISVITTGVSALRDKSHEDRQNRPNQKTSNEQHQ 47
RESULT 3
AAG82688
ID AAG82688 standard; Protein; 203 AA.
XX
AC AAG82688;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame protein sequence SEQ ID NO:2470.
XX
KM Staphylococcus epidermidis SRI strain; infection; diagnosis;
XX vaccination; endocarditis.
XX
OS Staphylococcus epidermidis.
XX
PN MO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000MO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI kimerly WJ;
XX
DR WPI: 2001-316495/33.
DR N-PSDB; AAH53538.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 18; Page 65L; 218pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the

CC	specification, but was obtained in electronic format directly from the
CC	USPTO web site.
XX	
SQ	Sequence 203 AA;
Query Match	28.3%; Score 58.5; DB 23; Length 203;
Best Local Similarity	48.3%; Pred. NO.4.1;
Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;	
OY	1 MNRLLEFGWIVLVSVLLGVASHIDNY 29 :::
Db	5 MNRLITLIG-VLLIILATLYFSKPIIDNY 32
RESULT 5	
AAG09408	
ID	AAG09408 standard; Protein: 291 AA.
XX	
AC	AAG09408;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 7330.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
RN	termination sequence.
XX	
CS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
EF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	28-APR-1999; 99US-0130891.
PR	30-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136392.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145952.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160818.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity

28.0%; Score 58; DB 21; Length 291;
34.5%; Pred. No. 7.3;

Matches 10; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

QY 3 RLIEHGWIVVSVLLGVASHIDNYP 31
||:| |:::|:| | : : |
DB 212 RLILLGLIMLIMGVLLVLTSSRLIRHP 240

RESULT 6
AAG09407
ID AAG09407 standard; Protein; 344 AA.
XX
AC AAG09407;
XX
XX 17-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana protein fragment SEQ ID NO: 7329.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143524.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157453.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158869.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161892.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 28.0% Score 58; DB 21; Length 344;

Best Local Similarity 34.5%; Pred. No. 8.9; Mismatches 0; Gaps 0;

Matches 10; Conservative 10; Indels 0; Gaps 0;

QY 3 RLIELGTMIIVVSVILLGVASHIDNYP 31
DB 265 RLILILGMLIMLGVLLVTSRLIRHLP 293

RESULT 7
AAG09406
ID AAG09406 standard; Protein; 361 AA.
XX
AC AAG09406;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 7328.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termindation sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PF 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.
XX
PR 23-APR-1999; 99US-0130891.
XX
PR 28-APR-1999; 99US-0131449.
XX
PR 30-APR-1999; 99US-0132048.
XX
PR 04-MAY-1999; 99US-0132484.
XX
PR 05-MAY-1999; 99US-0132485.
XX
PR 06-MAY-1999; 99US-0132486.
XX
PR 06-MAY-1999; 99US-0132487.
XX
PR 07-MAY-1999; 99US-0132863.
XX
PR 11-MAY-1999; 99US-0134256.
XX
PR 14-MAY-1999; 99US-0134218.
XX
PR 14-MAY-1999; 99US-0134221.
XX
PR 14-MAY-1999; 99US-0134221.
XX
PR 14-MAY-1999; 99US-0134370.
XX
PR 18-MAY-1999; 99US-0134768.
XX
PR 19-MAY-1999; 99US-0134941.
XX
PR 20-MAY-1999; 99US-0135124.
XX
PR 21-MAY-1999; 99US-0135353.
XX
PR 24-MAY-1999; 99US-0135629.
XX
PR 25-MAY-1999; 99US-0136021.
XX
PR 27-MAY-1999; 99US-0136592.
XX
PR 28-MAY-1999; 99US-0136782.
XX
PR 01-JUN-1999; 99US-0137222.
XX
PR 03-JUN-1999; 99US-0137528.
XX
PR 04-JUN-1999; 99US-0137502.
XX
PR 07-JUN-1999; 99US-0137724.
XX
PR 08-JUN-1999; 99US-0138094.
XX
PR 10-JUN-1999; 99US-0138540.
XX
PR 10-JUN-1999; 99US-0138847.
XX
PR 14-JUN-1999; 99US-0139119.
XX
PR 16-JUN-1999; 99US-0139452.
XX
PR 17-JUN-1999; 99US-0139453.
XX
PR 18-JUN-1999; 99US-0139454.
XX
PR 18-JUN-1999; 99US-0139454.
XX
PR 18-JUN-1999; 99US-0139455.
XX
PR 18-JUN-1999; 99US-0139456.
XX
PR 18-JUN-1999; 99US-0139457.
XX
PR 18-JUN-1999; 99US-0139458.

Query Match	Best Local Similarity	Matches	10; Conservative	Score 58.0%; Pred. No. 10; Mismatch
20-AUG-1999;	9905-0149929;			
PR 23-AUG-1999;	9905-0149930;			
PR 23-AUG-1999;	9905-0149930;			
PR 25-AUG-1999;	9905-0150566;			
PR 26-AUG-1999;	9905-0150984;			
PR 27-AUG-1999;	9905-0151065;			
PR 27-AUG-1999;	9905-0151066;			
PR 30-AUG-1999;	9905-0151080;			
PR 31-AUG-1999;	9905-0151133;			
PR 01-SEP-1999;	9905-0151330;			
PR 07-SEP-1999;	9905-0152673;			
PR 10-SEP-1999;	9905-0153070;			
PR 13-SEP-1999;	9905-0153758;			
PR 15-SEP-1999;	9905-0154018;			
PR 16-SEP-1999;	9905-0154039;			
PR 20-SEP-1999;	9905-0154779;			
PR 22-SEP-1999;	9905-0155139;			
PR 23-SEP-1999;	9905-0155586;			
PR 24-SEP-1999;	9905-0155659;			
PR 28-SEP-1999;	9905-0156458;			
PR 29-SEP-1999;	9905-0156596;			
PR 04-OCT-1999;	9905-0157117;			
PR 05-OCT-1999;	9905-0157753;			
PR 06-OCT-1999;	9905-0157865;			
PR 07-OCT-1999;	9905-0158029;			
PR 08-OCT-1999;	9905-0158232;			
PR 12-OCT-1999;	9905-0158593;			
PR 13-OCT-1999;	9905-0159294;			
PR 13-OCT-1999;	9905-0159295;			
PR 14-OCT-1999;	9905-0159329;			
PR 14-OCT-1999;	9905-0159330;			
PR 14-OCT-1999;	9905-0159637;			
PR 14-OCT-1999;	9905-0159638;			
PR 18-OCT-1999;	9905-0159584;			
PR 21-OCT-1999;	9905-0160741;			
PR 21-OCT-1999;	9905-0160767;			
PR 21-OCT-1999;	9905-0160768;			
PR 21-OCT-1999;	9905-0160770;			
PR 21-OCT-1999;	9905-0160814;			
PR 21-OCT-1999;	9905-0160815;			
PR 22-OCT-1999;	9905-0160980;			
PR 22-OCT-1999;	9905-0160981;			
PR 22-OCT-1999;	9905-0160989;			
PR 23-OCT-1999;	9905-0161044;			
PR 25-OCT-1999;	9905-0161405;			
PR 25-OCT-1999;	9905-0161406;			
PR 25-OCT-1999;	9905-0161359;			
PR 26-OCT-1999;	9905-0161360;			
PR 26-OCT-1999;	9905-0161361;			
PR 28-OCT-1999;	9905-0161920;			
PR 28-OCT-1999;	9905-0161922;			
PR 28-OCT-1999;	9905-0161993;			
PR 29-OCT-1999;	9905-0162142;			

XX DE Propionibacterium acnes immunogenic protein #14035.
 XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 XX KW dermatological; osteopathic; neuroprotectant.
 XX OS Propionibacterium acnes.
 XX PN WO200181581-A2.
 XX PD 01-NOV-2001.
 XX PF 20-APR-2001; 2001WO-US12865.
 XX PR 21-APR-2000; 2000US-199047P.
 XX PR 02-JUN-2000; 2000US-208841P.
 XX PR 07-JUL-2000; 2000US-216747P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX DR WPI: 2001-616774/71.
 XX DR N-PESDB: AAS59558.
 XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
 XX PT vaccinating against and diagnosing infections, especially useful for
 XX PT treating acne vulgaris -
 XX PS Example 1; SEQ ID No 14334; 1069pp; English.
 XX CC Sequences AU039105-AU06017 represent Propionibacterium acnes immunogenic
 XX CC polypeptides. The proteins and their associated DNA sequences are used in
 XX CC the treatment, prevention and diagnosis of medical conditions caused by
 XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 XX CC P. acnes is also involved in infections of bone, joints and the central
 XX CC nervous system, however it is particularly involved in the inflammatory
 XX CC lesions associated with acne vulgaris. A method for detecting the
 XX CC presence or absence of P. acnes in a patient comprises contacting a
 XX CC sample with a binding agent that binds to the proteins of the invention
 XX CC and determining the amount of bound protein in the sample. The
 XX CC polypeptides may be used as antigens in the production of antibodies
 XX CC specific for P. acnes proteins. These antibodies can be used to
 XX CC downregulate expression and activity of P. acnes polypeptides and
 XX CC therefore treat P. acnes infections. The antibodies may also be used as
 XX CC diagnostic agents for determining P. acnes presence, for example, by
 XX CC enzyme linked immunosorbent assay (ELISA).
 XX CC Note: The sequence data for this patent did not form part of the printed
 XX CC specification, but was obtained in electronic format directly from WIPO
 XX CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 327 AA:
 XX
 XX Query Match 27.5%; Score 57; DB 22; Length 327;
 XX Best Local Similarity 45.5%; Pred. No. 12;
 XX Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 QY 7 LTGMIVLVSVLLGVASHIDN 28
 DB 249 LAGMIVLVGGLGAVAAHYVN 270
 RESULT 9
 ABB54920
 XX ID ABB54920 standard; Protein: 320 AA.
 XX AC ABB54920;
 XX DT 16-MAY-2002 (first entry)

XX DE Lactococcus lactis protein yqbf.
 XX KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX OS Lactococcus lactis IL1403.
 XX PN FR2807446-A1.
 XX PD 12-OCT-2001.
 XX PF 11-APR-2000; 2000FR-0004630.
 XX PR 11-APR-2000; 2000FR-0004630.
 XX PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX PI Bolotine A, Sorokline A, Renault P, Ehrlich SD;
 XX PI WPI: 2002-043418/06.
 XX PT New nucleotide sequence useful in the identification or Lactococcus
 XX PT lactis and related species -
 XX PS Claim 6; SEQ ID No 1622; 2504pp; French.
 XX CC The present invention is related to a Lactococcus lactis nucleotide
 XX CC sequence (ABN90521) and related proteins (ABB53300-ABB55621). The
 XX CC nucleic acid sequence is useful in the detection and/or amplification of
 XX CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 XX CC related species. The proteins of the invention are useful for the
 XX CC biosynthesis or biodegradation of a composition of interest. The
 XX CC invention helps research in lactic bacteria, particularly useful in the
 XX CC production of yogurt and cheese.
 XX CC Note: The sequence data for this patent is based on equivalent patent
 XX CC WO200177334 (published 18-OCT-2001) which is available in electronic
 XX CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 320 AA:
 XX
 XX Query Match 26.6%; Score 55; DB 23; Length 320;
 XX Best Local Similarity 37.9%; Pred. No. 21;
 XX Matches 11; Conservative 8; Mismatches 10; Indels 0; Gaps 0;
 QY 1 MNRLIELTGMIVLVSVLLGVASHIDN 29
 DB 1 MKRRKMIAMILVLSVLVGGYFAIGNY 29
 RESULT 10
 AAG92354
 XX ID AAG92354 standard; Protein: 900 AA.
 XX AC AAG92354;
 XX DT 26-SEP-2001 (first entry)
 XX DE C glutamicum protein fragment SEQ ID NO: 6108.
 XX KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
 XX KW organic acid synthesis.
 XX OS Corynebacterium glutamicum.
 XX PN EP1108790-A2.
 XX PD 20-JUN-2001.
 XX PF 18-DEC-2000; 2000EP-0127688.
 XX PR 16-DEC-1999; 99JP-0377484.
 XX PR 07-APR-2000; 2000JP-0159162.
 XX PR 03-AUG-2000; 2000JP-0280988.

```

XX      (K10W ) KIYOMA HAKKO KOGYO KK.
PA
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI: 2001-376931/40.
DR N-PSDB: AAH67573.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
XX Claim 17: SEQ ID NO: 6108; 246bp + Sequence listing; English.
PS
XX
XX The present invention provides a number of nucleotide and protein
XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of coryneform bacterium, measuring expression amount and
XX analysing the expression profile or expression pattern of a gene derived
XX from Coryneform bacterium, and identifying a homologue of a gene derived
XX from coryneform bacterium. Coryneform bacteria are useful for producing
XX amino acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a protein described
XX in the exemplification of the invention.
XX
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from the
XX European Patent Office.
XX
XX Sequence 900 AA:
SQ
XX
XX Query Match 26.3%; Score 54.5; DB 22; Length 900;
XX Best Local Similarity 34.2%; Pred. No. 84;
XX Matches 13; Conservative 9; Mismatches 13; Indels 3; Gaps 1;
XX
XX QY 6 ELTG---WIVLVSVIILGVASHIDNYQPPEDSASVQH 40
XX ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
XX Db 625 ELTDADRWIVDLEQVRIDVDAYILDNYQFAKAMEELXH 662
XX
XX RESULT 11
XX ABG20939
XX ID ABG20939 standard; Protein: 1048 AA.
XX
XX AC ABG20939;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #20930.
XX
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX PN 11-OCT-2001.
XX
XX PD 30-MAR-2001; 2001WO-US08631.
XX
XX PF 31-MAR-2000; 2000US-0540217.
XX
XX PR 23-AUG-2000; 2000US-0649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX
XX DR N-PSDB: AAS85126.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in

```

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID NO 51298; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPRO
CC at ftp.wipro.int/pub/published_pct_sequences.
XX
XX Sequence 1048 AA;
SQ

Query Match 26.1%; Score 54; DB 22; Length 1048;
Best Local Similarity 38.2%; Pred. NO. 1.2e+02;
Matches 13; Conservative 9; Mismatches 8; Indels 4; Gaps 2;

OY 11 IYLVSVLLG--VASHIDNYQPE-QSASYQH 40
:|::||::||| | : |||::: |
Db 39 VVITSMWLLGRSIQASRKEMQPEKETPEVLH 72

RESULT 12
AABL1050
ID AABL1050 standard; Protein: 206 AA.
XX
XX AABL1050;
AC
XX
DT 16-FEB-2001 (first entry)
DE Staphylococcus aureus sortase transamidase srtA protein.
XX
XX Staphylococcus aureus sortase transamidase srtA protein.
KW Sortase transamidase; Gram-positive bacterium; covalent cross-link;
RV peptidoglycan; antibacterial; vaccine; treatment; detection; srtA;
KW bacterial infection.
XX
XX Staphylococcus aureus.
CS
SN WO200062804-A2.
PN
XX
PD 26-OCT-2000.
PF 13-APR-2000; 2000WO-US10196.
XX
PR 15-APR-1999; 99US-0292437.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Schneewind O, Mazmanian S, Liu G, Ton-That H;
DR WPI: 2000-665197/64.
XX N-PSTDB: AAC65625.
XX
PT A substantially purified sortase-transamidase from a Gram-positive
PT bacterium for use in the treatment and detection of Gram-positive
PT bacterial infections -


```

CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
CC
XX
XX
Sequence 455 AA;
50
Query Match 25.1%; Score 52; DB 22; Length 455;
Best Local Similarity 32.6%; Pred. No. 85;
Matches 14; Conservative 8; Mismatches 13; Indels 8; Gaps 2
QY 6 ELTGW-----IVLVSVTLIGVASHIIDNQPEQ---SASVQH 40
|:| |:::|::| |::| |
Db 163 EMEFWFAMIKIVTITALLVGGYLAITGFGPPRGVPAFSFH 205

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 21:37:40 ; Search time 15 Seconds
(without alignments)
80.423 Million cell updates/sec

Title: US-09-828-456-2

Perfect score: 207
Sequence: 1 MNRLELTGMIVLVSVILL.....VASHIDNTPPEOSASVQHK 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.5	29.2	231	4	US-09-134-001C-3597 Sequence 3597, Ap
2	58.5	28.3	203	4	US-09-134-001C-4445 Sequence 4445, Ap
3	52	25.1	285	1	US-08-292-045-10 Sequence 10, Appl
4	50.5	24.4	491	4	US-09-134-001C-3510 Sequence 3510, Ap
5	50	24.2	265	1	US-08-292-045-2 Sequence 2, Appl1
6	48.5	23.4	575	3	US-08-222-865-2 Sequence 2, Appl1
7	48.5	23.4	575	3	US-09-510-949-2 Sequence 2, Appl1
8	46	22.2	309	4	US-09-134-001C-5667 Sequence 5667, Ap
9	46	22.2	548	4	US-09-601-091-2 Sequence 2, Appl1
10	46	22.2	548	4	US-09-601-091-2 Sequence 2, Appl1
11	46	22.2	548	4	US-09-398-395A-52 Sequence 52, Appl
12	45.5	22.0	249	4	US-09-134-001C-4442 Sequence 4442, Ap
13	45.5	22.0	362	3	US-09-120-365-74 Sequence 74, Appl
14	45.5	22.0	456	4	US-09-515-039-74 Sequence 74, Appl
15	45.5	22.0	642	1	US-09-134-001C-3771 Sequence 3771, Ap
16	45.5	22.0	642	1	US-08-706-936-2 Sequence 2, Appl1
17	45.5	22.0	643	2	US-08-616-844-39 Sequence 39, Appl
18	45.5	22.0	643	2	US-08-599-654-39 Sequence 39, Appl
19	45.5	22.0	643	3	US-08-944-868A-39 Sequence 39, Appl
20	45.5	22.0	643	3	US-08-944-868A-39 Sequence 39, Appl
21	45.5	22.0	643	3	US-08-944-423A-39 Sequence 39, Appl
22	45	21.7	80	3	US-09-081-320-31 Sequence 31, Appl
23	45	21.7	80	4	US-09-574-141A-31 Sequence 31, Appl
24	45	21.7	80	4	US-09-707-780-31 Sequence 31, Appl
25	45	21.7	236	4	US-09-134-001C-5548 Sequence 5548, Ap
26	45	21.7	691	1	US-08-178-477B-2 Sequence 2, Appl1
27	44.5	21.5	142	3	US-08-908-643C-53 Sequence 53, Appl

28	44.5	21.5	166	3	US-08-908-643C-51	Sequence 51, Appl
29	44.5	21.5	287	3	US-08-549-515-10	Sequence 10, Appl
30	44.5	21.5	322	4	US-09-457-046B-14	Sequence 14, Appl
31	44.5	21.5	445	4	US-09-457-046B-52	Sequence 52, Appl
32	44.5	21.5	572	2	US-08-467-963C-4	Sequence 4, Appl1
33	44.5	21.5	572	2	US-08-838-189D-4	Sequence 4, Appl1
34	44.5	21.5	572	3	US-08-852-344D-4	Sequence 4, Appl1
35	44.5	21.5	572	3	US-08-344-639E-4	Sequence 4, Appl1
36	44.5	21.5	572	3	US-08-987-439-3	Sequence 4, Appl1
37	44.5	21.5	572	4	US-08-467-969A-4	Sequence 4, Appl1
38	44.5	21.5	572	4	US-08-467-969A-4	Sequence 4, Appl1
39	44.5	21.5	572	4	US-08-001-554A-4	Sequence 4, Appl1
40	44	21.3	239	4	US-09-134-001C-4165	Sequence 4165, Ap
41	44	21.3	279	4	US-09-134-001C-4900	Sequence 4900, Ap
42	44	21.3	292	4	US-09-134-001C-4109	Sequence 4109, Ap
43	44	21.3	458	4	US-08-487-596-16	Sequence 16, Appl
44	44	21.3	489	2	US-08-663-566A-6	Sequence 6, Appl1
45	44	21.3	489	2	US-08-023-610-6	Sequence 6, Appl1

ALIGNMENTS

```
RESULT 1
US-09-134-001C-3597
Sequence 3597, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3597
LENGTH: 231
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3597

Query Match 29.2%; Score 60.5; DB 4; Length 231;
Best Local Similarity 35.0%; Pred. No. 0.31;
Matches 14; Conservative 12; Mismatches 7; Indels 7; Gaps 2;

QY 9 GWIVLVSVILLGVA-----SHD--NYRPEOSASVQHK 41
      | | | | | | | | | | | | | | | | | | | | |
Db 8 GIIIFITVITIGVSLRDKSHEDRONRPNOKTSNBOHQ 47

RESULT 2
US-09-134-001C-4445
Sequence 4445, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4445
LENGTH: 203
TYPE: PRT
```

ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4445

Query Match 28.3%; Score 58.5; DB 4; Length 203;
Best Local Similarity 48.3%; Pred. No. 0.52;
Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 1 MNRLIEITGWIYVSVVILLGVASHIDNY 29
||||| : : : : :
DB 5 MNRLITLIG-VLLITLITLIFSKPYIDNY 32

RESULT 3
US-08-292-045-10
Sequence 10, Application US/08292045

PATENT No. 5602005
GENERAL INFORMATION:
APPLICANT: HERR, JOHN C.
TITLE OF INVENTION: PRIMATE INTRA-ACROSOMAL SPERM ANTIGEN
TITLE OF INVENTION: FOR USE IN A CONTRACEPTIVE VACCINE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,045
FILING DATE:
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/318,551
FILING DATE: 03-MAR-1989
APPLICATION NUMBER: US 07/481,491
FILING DATE: 16-FEB-1990
APPLICATION NUMBER: US 07/858,798
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5602005man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 494-109-0 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-292-045-10

Query Match 25.1%; Score 52; DB 1; Length 285;
Best Local Similarity 36.6%; Pred. No. 7.1;
Matches 15; Conservative 7; Mismatches 11; Indels 8; Gaps 2;

QY 1 MNRLIEITGWIYVSVVILLGVASHIDNYOPPEQASVQHK 41
||| : : : : :
DB 1 MNRL-----LLMSLYLLGSANGTSS-QPNESGSDIDH 33

RESULT 4
US-09-134-001C-3510

Sequence 3510, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3510
LENGTH: 491
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3510

Query Match 24.4%; Score 50.5; DB 4; Length 491;
Best Local Similarity 31.7%; Pred. No. 23;
Matches 13; Conservative 9; Mismatches 8; Indels 11; Gaps 2;

QY 12 VLVSVYLLGVAS-----HIDNTQPEQASV---VQHK 41
::||: : : : :
DB 29 IIVASMLFIGGSVHAESNHLNQSEVSASHSTEMQHK 69

RESULT 5
US-08-292-045-2
Sequence 2, Application US/08292045

PATENT No. 5602005
GENERAL INFORMATION:
APPLICANT: HERR, JOHN C.
TITLE OF INVENTION: PRIMATE INTRA-ACROSOMAL SPERM ANTIGEN
TITLE OF INVENTION: FOR USE IN A CONTRACEPTIVE VACCINE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,045
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/318,551
FILING DATE: 03-MAR-1989
APPLICATION NUMBER: US 07/481,491
FILING DATE: 16-FEB-1990
APPLICATION NUMBER: US 07/858,798
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5602005man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 494-109-0 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

Query Match 25.1%; Score 52; DB 1; Length 285;
Best Local Similarity 36.6%; Pred. No. 7.1;
Matches 15; Conservative 7; Mismatches 11; Indels 8; Gaps 2;

QY 1 MNRLIEITGWIYVSVVILLGVASHIDNYOPPEQASVQHK 41
||| : : : : :
DB 1 MNRL-----LLMSLYLLGSANGTSS-QPNESGSDIDH 33

RESULT 4
US-09-134-001C-3510


```
RESULT 10
US-09-601-091-4
; Sequence 4, Application US/09601091
; Patent No. 6342380
; GENERAL INFORMATION:
; APPLICANT: Colby, S. M. et al.
; TITLE OF INVENTION: Germacrene C Synthase Gene of Lycopersicon Esculentum
; FILE REFERENCE: 4630-55758
; CURRENT APPLICATION NUMBER: US/09/601,091
; CURRENT FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: PCT/US99/02133
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,579
; PRIOR FILING DATE: 1998-02-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-601-091-4

Query Match          22.2%; Score 46; DB 4; Length 548;
Best Local Similarity 40.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

QY 11 IVLVSVILGVASHIDNYQPEQSASVQH 40
    |||:: |||| | | | | | | | | | |
Db 61 LVLIDAMQRLGVAYHFBN----ELETISION 86

RESULT 11
US-09-398-395A-52
; Sequence 52, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6468772, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-398-395A-52

Query Match          22.2%; Score 46; DB 4; Length 548;
Best Local Similarity 40.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

QY 11 IVLVSVILGVASHIDNYQPEQSASVQH 40
    |||:: |||| | | | | | | | | | |
Db 61 LVLIDAMQRLGVAYHFBN----ELETISION 86

RESULT 12
US-09-134-001C-4442
; Sequence 4442, Application US/09134001C
; Patent No. 6380370
```

```
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4442
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4442

Query Match          22.0%; Score 45.5; DB 4; Length 249;
Best Local Similarity 31.4%; Pred. No. 54;
Matches 11; Conservative 7; Mismatches 10; Indels 7; Gaps 1;

QY 8 TGMT-----VLVSVILGVASHIDNYQPEQS 35
    |||: ||:: | | | | | | | | | |
Db 66 TGMELVLSITIFVLLIALLYGAYVYHYISDEQS 100

RESULT 13
US-09-120-365-74
; Sequence 74, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
; APPLICANT: Natoli, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/120,365
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 74
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Vigna mungo CP
US-09-120-365-74

Query Match          22.0%; Score 45.5; DB 3; Length 362;
Best Local Similarity 34.3%; Pred. No. 86;
Matches 12; Conservative 10; Mismatches 8; Indels 5; Gaps 2;

QY 1 MNRLEITGMVILVSVILGVASHIDNYQPEQS 35
    |:: |:: ||:: ||:: ||:: | | | | | |
Db 3 MKKLL---WVVLSTSLV-LGVANSFDFHEKDLKS 32

RESULT 14
US-09-515-039-74
; Sequence 74, Application US/09515039
; Patent No. 6214599
; GENERAL INFORMATION:
; APPLICANT: Natoli, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/515,039
; CURRENT FILING DATE: 2000-03-06
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 74
; LENGTH: 362
; TYPE: PRT
```

ORGANISM: *Vigna mungo* CP
US-09-515-039-74

Query Match	22.08;	Score 45.5;	DB 4;	Length 362;
Best Local Similarity	34.38;	Pred. No. 86;		
Matches 12; Conservative	10;	Mismatches 8;	Indels 5;	Gaps 2

```
QY      1 MNRLIELTGWIVLVSVILLGVASHIDNQPPEQS   35
        | :| :|| :||| : ||| : | :: :|
Db      3 MKKLL---WVVLSTSLV-LGVANSFDFHEKDLES  32
```

```

1 RESULT 15
2 US-09-134-001C-3771
3 : Sequence 3771, Application US/09134001C
4 : Patent No. 6380370
5 : GENERAL INFORMATION:
6 : APPLICANT: Lynn Doucette-Stamm et al
7 : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
8 : TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
9 : FILE REFERENCE: GTC-007
10 : CURRENT APPLICATION NUMBER: US/09/134,001C
11 : CURRENT FILING DATE: 1998-08-13
12 : PRIOR APPLICATION NUMBER: US 60/064,964
13 : PRIOR FILING DATE: 1997-11-08
14 : PRIOR APPLICATION NUMBER: US 60/055,779
15 : PRIOR FILING DATE: 1997-08-14
16 : NUMBER OF SEQ ID NOS: 5674
17 : SEQ ID NO 3771
18 : LENGTH: 456
19 : TYPE: PRT
20 : ORGANISM: Staphylococcus epidermidis
21 US-09-134-001C-3771

```

Query Match	22.0%;	Score 45.5;	DB 4;	Length 456;
Best Local Similarity	27.5%;	Pred. No. 1.1e+02;		
Matches 11;	Conservative 11;	Mismatches 13;	Indels 5;	Gaps 1

```

Qy      6 ELTGW-----IYLVSVALLGVASHINDYQPEOSASVQH 40
        || |      || :::::|| : :: | : |
Db     149 ELEFWPAIKIYVILLALVIGVMILFAKTPRGNTSLTH 188

```

Search completed: February 6, 2003, 21:39:54
Job time : 17 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 21:38:20 : Search time 11 Seconds

(without alignments)
82.631 Million cell updates/sec

Title: US-09-828-456-2

Perfect score: 207

Sequence: 1 MNRLELTGMIVLVSVILL.....VASHIDNYPQASVQHK 41

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 129505 seqs, 22169297 residues

Total number of hits satisfying chosen parameters: 129505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications-AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	207	100.0	41	10	US-09-828-456-2
2	54.5	26.3	900	9	US-09-738-626-6108
3	53.5	25.8	206	9	US-09-933-999A-3
4	53.5	25.8	206	10	US-09-815-242-5751
5	53.5	25.8	206	10	US-09-815-242-12676
6	49	23.7	1053	10	US-09-815-242-5136
7	48.5	23.4	323	9	US-09-978-295A-330
8	48.5	23.4	323	9	US-09-978-697-330
9	48.5	23.4	323	9	US-09-978-192A-330
10	48.5	23.4	323	9	US-09-999-832A-330
11	48.5	23.4	323	9	US-09-978-189-330
12	48.5	23.4	323	9	US-10-028-072-272
13	48.5	23.4	323	9	US-10-121-049-272
14	48.5	23.4	323	9	US-10-123-904-272
15	48.5	23.4	323	9	US-10-140-470-272
16	48.5	23.4	360	10	US-09-815-242-12011
17	48	23.2	81	9	US-09-738-626-5932
18	48	23.2	497	12	US-10-033-109-10
19	47	22.7	178	9	US-09-738-626-5509

20	47	22.7	397	10	US-09-815-242-12699	Sequence 12699, A
21	46.5	22.5	83	10	US-09-864-761-46530	Sequence 46530, A
22	46.5	22.5	444	10	US-09-881-752A-212	Sequence 212, App
23	46	22.2	530	9	US-10-047-542-85	Sequence 85, App1
24	46	22.2	548	9	US-09-738-626-5776	Sequence 5776, App
25	46	22.2	548	10	US-09-887-586A-52	Sequence 52, App1
26	46	22.2	548	10	US-09-903-012-52	Sequence 77, App1
27	45.5	22.0	44	10	US-09-739-907-77	Sequence 160, App
28	45.5	22.0	58	10	US-09-739-907-160	Sequence 639, App
29	45.5	22.0	61	10	US-09-764-847-639	Sequence 74, App1
30	45.5	22.0	537	9	US-10-047-542-74	Sequence 11, App1
31	45.5	22.0	643	9	US-10-060-763-11	Sequence 39, App1
32	45.5	22.0	643	10	US-09-371-900-39	Sequence 56, App1
33	45.5	22.0	643	10	US-09-924-417-56	Sequence 11, App1
34	45.5	22.0	643	12	US-10-063-763-11	Sequence 55, App1
35	45	21.7	94	10	US-09-778-927A-55	Sequence 1330, App
36	45	21.7	154	10	US-09-867-550-1330	Sequence 6459, App
37	45	21.7	159	9	US-09-738-626-6459	Sequence 5178, App
38	45	21.7	162	9	US-09-738-626-5178	Sequence 43, App1
39	45	21.7	299	10	US-09-510-332-9	Sequence 9, App1
40	45	21.7	299	10	US-09-393-634-43	Sequence 6415, App
41	45	21.7	317	9	US-09-738-626-6415	Sequence 6555, App
42	45	21.7	449	9	US-09-738-626-6555	Sequence 28, App1
43	45	21.7	464	9	US-10-024-623-28	Sequence 39, App1
44	45	21.7	473	10	US-09-790-264-39	Sequence 84, App1
45	45	21.7	535	9	US-10-047-542-84	

ALIGNMENTS

RESULT 1
US-09-828-456-2
Sequence 2, Application US/09828456
Patent No. US20020051982A1
GENERAL INFORMATION:
APPLICANT: Levy, Stewart et al.
TITLE OF INVENTION: NO. US20020051982A1 BLR Molecules Affecting Antibiotic Susce
FILE REFERENCE: PKZ-030
CURRENT APPLICATION NUMBER: US/09/828,456
CURRENT FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/195,505
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/218,380
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 41
TYPE: PRT
ORGANISM: Escherichia coli
US-09-828-456-2

Query Match 100.0%; Score 207; DB 10; Length 41;
Best Local Similarity 100.0%; Pred No. 1.5e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLELTGMIVLVSVILLGVASHIDNYPQASVQHK 41
DB 1 MNRLELTGMIVLVSVILLGVASHIDNYPQASVQHK 41

RESULT 2
US-09-738-626-6108
Sequence 6108, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO

```

APPLICANT: TATEISHI, NAOAO
APPLICANT: SENOH, AKIHITO
APPLICANT: IKEDA, MASATO
APPLICANT: OKAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: Patentin ver. 3.0
SEQ ID NO 6108
LENGTH: 900
TYPE: PRF
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6108

```

Query Match	26.3%	Score	54.5	DB	9	Length	900
Best Local Similarity	34.2%	Pred. No.	18				
Matches	13	Conservative	9	Mismatches	13	Indels	3
						Gaps	1

QY	6	ELTG---WIVLVSVILLGVASHIDNQPEQSASYH	40
		: : :	
Db	625	ELTDADRNIVDLQVRIDYDAYLDNQFAKANEELVH	662

```

RESULT 3
US-09-933-999A-3
: Sequence 3, Application US/09933999A
: Publication No. US20030022178A1
: GENERAL INFORMATION:
: APPLICANT: U.C. Regents
: APPLICANT: Schneewind et al.
: TITLE OF INVENTION: Identification of Sortase Genes
: FILE REFERENCE: 510015-249
: CURRENT APPLICATION NUMBER: US/09/933,999A
: CURRENT FILING DATE: 2001-12-20
: NUMBER OF SEQ ID NOS: 40
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 206
: TYPE: PRT
: ORGANISM: Staphylococcus aureus
US-09-933-999A-3

```

Query Match	25.8%	Score	53.5	DB	9	Length	206
Best Local Similarity	42.9%	Pred.	No. 4.4				
Matches	12	Conservative	5	Mismatches	10	Indels	1
						Gaps	1

QY	2	NRLEIETGIVLVSVILLGVASHIDNY	29
Db	6	NRLEMTIAG-VVLLIVAAVLEAKPHIDNY	32

```

RESULT 4
US-09-815-242-5751
; Sequence 5751, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in

```

```

1  TITLE OF INVENTION:  Prokaryotes
2  FILE REFERENCE:  ELITRA-011A
3  CURRENT APPLICATION NUMBER:  US/09/815,242
4  CURRENT FILING DATE:  2001-03-21
5  PRIOR APPLICATION NUMBER:  60/191,078
6  PRIOR FILING DATE:  2000-03-21
7  PRIOR APPLICATION NUMBER:  60/206,848
8  PRIOR FILING DATE:  2000-05-23
9  PRIOR APPLICATION NUMBER:  60/207,727
10 PRIOR FILING DATE:  2000-05-26
11 PRIOR APPLICATION NUMBER:  60/242,578
12 PRIOR FILING DATE:  2000-10-23
13 PRIOR APPLICATION NUMBER:  60/253,625
14 PRIOR FILING DATE:  2000-11-27
15 PRIOR APPLICATION NUMBER:  60/257,931
16 PRIOR FILING DATE:  2000-12-22
17 PRIOR APPLICATION NUMBER:  60/269,308
18 PRIOR FILING DATE:  2001-02-16
19 NUMBER OF SEQ ID NOS:  14110
20 SOFTWARE:  FastSeq for Windows Version 4.0
21 SEQ ID NO 5751
22 LENGTH:  206
23 TYPE:  prt
24 ORGANISM:  Staphylococcus aureus
25 US-09-815-242-5751

```

Query Match	25.8%	Score 53.5	DB 10	length 206
Best Local Similarity	42.9%	Pred. No. 4.4		
Matches 12, Conservative	5	Mismatches 10	Indels 1	Gaps 1

Qy	2	NRIELGTGIVLVSVILGVASHIDNY	29
		: :	
Db	6	NRLMTIAG-VVILVAAYLFAKPHIDNY	32

```

RESULT 5
US-09-815-242-12676
: Sequence 12676, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in

```

```

1  FILE REFERENCE: ELITRA.011A
2  CURRENT APPLICATION NUMBER: US/09/815,242
3  CURRENT FILING DATE: 2001-03-21 070

```

1 PRIORITY FILING DATE: 2000-03-21
2
3 PRIORITY APPLICATION NUMBER: 60/206,848
4
5 PRIORITY FILING DATE: 2000-05-23
6

```

1 PRIOR APPLICATION NUMBER: 60/242,578
2 PRIOR FILING DATE: 2000-10-23
3 PRIOR APPLICATION NUMBER: 60/253,625
4 PRIOR FILING DATE: 2000-11-27
5 PRIOR APPLICATION NUMBER: 60/257,931
6 PRIOR FILING DATE: 2000-12-22
7 PRIOR APPLICATION NUMBER: 60/269,308
8 PRIOR FILING DATE: 2001-02-16
9 NUMBER OF SEQ ID NOS: 14110
10 SOFTWARE: FastSeq for Windows Version 4.0.0
11 SEQ ID NO 12676
12 LENGTH: 206
13 TYPE: prt
14 ORGANISM: Staphylococcus aureus

```


APPLICANT: Hillan, Kenneth J
APPLICANT: KJavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441

PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 23.4%; Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

Qy 8 TGMIV---LVVSVIL-----GVASHIDNYQPPQ 34
Db 238 SCWILTTTLVLSVVLWICCATVAVAEQYVPSK 273

RESULT 9
US-09-978-192A-330
Sequence 330, Application US/09978192A
Patent No. US2002017753A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Rong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann

APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334

PRIOR FILING DATE: 1998-04-01	PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08	PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08	PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08	PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08	PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09	PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09	PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21	PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21	PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23	PRIOR APPLICATION NUMBER: 60/083366
PRIOR FILING DATE: 1998-04-27	PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30	PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05	PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06	PRIOR APPLICATION NUMBER: 60/084411
PRIOR FILING DATE: 1998-05-06	PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084640

```

      Prior APPLICATION NUMBER: 60/084558
      Prior FILING DATE: 1998-05-07
      Prior APPLICATION NUMBER: 60/084600
      Prior FILING DATE: 1998-05-07
      Prior APPLICATION NUMBER: 60/084627
      Prior FILING DATE: 1998-05-07
      Prior APPLICATION NUMBER: 60/084643
      Prior FILING DATE: 1998-05-07
      Prior APPLICATION NUMBER: 60/085339
      Prior FILING DATE: 1998-05-13
      Prior APPLICATION NUMBER: 60/085338
      Prior FILING DATE: 1998-05-13
      Prior APPLICATION NUMBER: 60/085323
      Prior FILING DATE: 1998-05-13
      Prior APPLICATION NUMBER: 60/085862
      Prior FILING DATE: 1998-05-15
      Prior APPLICATION NUMBER: 60/085700
      Prior FILING DATE: 1998-05-15
      Prior APPLICATION NUMBER: 60/085689
      Prior FILING DATE: 1998-05-15
      Prior APPLICATION NUMBER: 60/085579
      Prior FILING DATE: 1998-05-15
      Prior APPLICATION NUMBER: 60/085580
      Prior FILING DATE: 1998-05-15
      Prior APPLICATION NUMBER: 60/085573
      Prior FILING DATE: 1998-05-15
      Prior APPLICATION NUMBER: 60/085704
      Prior FILING DATE: 1998-05-15
      Prior APPLICATION NUMBER: 60/085697

Query Match          23.4%; Score 48.5; DB 9; Length 323;
Best Local Similarly 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 9

Oy    8 TGWIV---LTVSVLL-----GVASHIDNPPEQ 34
Db     238 SGWLTTTLVLSVMVLMLCCATVAIVEQTVPSEK 273

           :|||: ||:||::||         ||:::||| |: 
US-09-999-832A-330                                US-09-999-832A-330
Sequence 330, Application US/09999832A
Publication No. US20020192706A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guiney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuio, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same
```

FILE REFERENCE: P2630PIC63
CURRENT APPLICATION NUMBER: US/09/999,832A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643

;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 23.4%; Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

OY 8 TGMIV--LVSVLL-----GVASHIDNYQPEQ 34
DB 238 SGWILTTTLVLSVWMLMICATVATAVEQYVPSEK 273

RESULT 11
US-09-978-189-330
; Sequence 330, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kiljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Peoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC7
; CURRENT APPLICATION NUMBER: US/09/978,189
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17

;; PRIOR APPLICATION NUMBER: 60/064249
;; PRIOR FILING DATE: 1997-11-03
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066364
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/077450
;; PRIOR FILING DATE: 1998-03-10
;; PRIOR APPLICATION NUMBER: 60/077632
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077641
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077649
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/07791
;; PRIOR FILING DATE: 1998-03-12
;; PRIOR APPLICATION NUMBER: 60/078004
;; PRIOR FILING DATE: 1998-03-13
;; PRIOR APPLICATION NUMBER: 60/078886
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078936
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955

PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13

PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 23.4% Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

OY 8 TGMIV---LVSVYIL-----GVASHIDNYPPO 34
DB 238 SGWILFTTLVSYMWLWICATVATVAGVYPSK 273

RESULT 12
US-10-028-072-272
; Sequence 272, Application US/10028072
; Publication No. US20030004311A1

GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang

TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24

[illegible]

Mon Feb 10 09:26:00 2003

us-09-828-456-2.rapb

Page 13

DB 238 SGWILFTTJLVLSYMLMTCATVATAVEGYVPSK 273

Search completed: February 6, 2003, 21:40:11
Job time : 12 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:05:32 ; Search time 2578 Seconds

(without alignments)
5429.965 Million cell updates/sec

Title: US-09-828-456-1

Perfect score: 481
Sequence: 1 tgcctcctatggaagatgca.....gtgatgtcagagatgtaa 481Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	481	100.0	481	6	AX278241	AX278241 Sequence
2	469	97.5	13172	1	AE000258	Escherich
3	469	97.5	15332	1	D90806	D90806 E.coli geno
4	469	97.5	15601	1	D90808	D90808 E.coli geno
5	469	97.5	20231	1	D90807	D90807 E.coli geno
6	452.2	94.0	13259	1	AE005386	Escherich
7	449.8	93.5	257071	1	AF002558	Escherich
8	347	72.1	358	1	AF219227	Escherich
9	203	42.2	19211	1	D90805	D90805 E.coli geno
10	106.6	22.2	23009	1	AE008763	Salmonell
11	103.4	21.5	233050	1	AL627271	Salmonell
12	42.4	8.8	134978	2	AC079004	Homo sapi
13	42.4	8.8	164184	2	AC079050	Homo sapi
14	42.4	8.8	167577	9	AC099687	Homo sapi
15	42.2	8.8	190015	2	AL591126	Mus muscu
16	41.8	8.7	67890	8	NCB14H13	Neurospor
17	41.2	8.6	175111	2	AC101999	Mus muscu
18	41.2	8.6	204591	2	AC122262	Mus muscu
19	41	8.5	95417	9	AC010876	Homo sapi
20	40.6	8.4	160824	2	AC069118	Homo sapi
21	40.6	8.4	173770	2	AC122716	Homo sapi
22	40.4	8.4	41118	2	AC090895	Homo sapi
23	40.2	8.4	106859	2	AC010295	Homo sapi
24	40	8.3	7448	3	AF465305	Dictyoste
25	40	8.3	18446	8	AB026660	Arabidops
26	40	8.3	19132	2	ATH131517	Arabidops
27	40	8.3	157672	2	AC100742	Mus muscu
28	40	8.3	178493	2	AL354981	Homo sapi
29	39.8	8.3	11321	1	AE006407	Lactococc
30	39.8	8.3	11341	1	AE007813	Clostridi
31	39.8	8.3	74589	8	AB010068	Arabidops
32	39.4	8.2	12668	3	AF079445	Dictyoste
33	39.4	8.2	58753	3	AF125459	Caenorhab
34	39.4	8.2	164051	9	AC106892	Homo sapi
35	39.2	8.1	145938	2	AC128165	Rattus no
36	39	8.1	4183	3	AF474377	Dictyoste
37	39	8.1	180854	9	AC107960	Homo sapi
38	38.8	8.1	197092	2	AC036155	Homo sapi
39	38.8	8.1	6294	6	AX344308	Sequence
40	38.8	8.1	9997	6	AX281373	Sequence
41	38.8	8.1	189235	2	AC113717	Rattus no
42	38.8	8.1	195411	2	AC092508	Mus muscu
43	38.6	8.0	85421	8	AB007651	Arabidops
44	38.6	8.0	149715	2	AC117165	Rattus no
45	38.6	8.0	149715	2	AC117165	Rattus no

ALIGNMENTS

RESULT 1
AX278241 481 bp
LOCUS AX278241
DEFINITION Sequence 1 from Patent WO0177176.
ACCESSION AX278241
VERSION AX278241.1 GI:16605287
KEYWORDS
SOURCE
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
1
Levy,S.B. and McMurry,L.M.
TITLE Novel b1r molecules affecting antibiotic susceptibility
JOURNAL Patent: WO 01/77176-A 1 18-OCT-2001.

TRUSTERS OF TUFTS COLLEGE (US)
 FEATURES Location/Qualifiers
 source 1. 481
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 CDS 239..364
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD10539.1"
 /db_xref="GI:16605288"
 /translation="MNRIELTGWIVLVVSYLLGVASHIDNPPEOSASVQHK"
 BASE COUNT 139 a 95 c 100 g 147 t
 ORIGIN
 Query Match 100.0%; Score 481; DB 6; Length 481;
 Best Local Similarity 100.0%; Pred. No. 2e-115;
 Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCCCTCATTTAGATGATGCTATTGACATATCTTACCCAGCTCAAGATT 60
 DB 1 TGCCCTCATTTAGATGATGCTATTGACATATCTTACCCAGCTCAAGATT 60
 QY 61 AATAACCTGCCAGCAATAAGGAGATGTTGTTTAAGTCAAAAAAATAGCGAATTTTC 120
 DB 61 AATAACCTGCCAGCAATAAGGAGATGTTGTTTAAGTCAAAAAAATAGCGAATTTTC 120
 QY 121 CAACGACAAAGCTAAATATCGCAAAAACCTCAGTAAAAATCTCTGAGCTATTAT 180
 DB 121 CAACGACAAAGCTAAATATCGCAAAAACCTCAGTAAAAATCTCTGAGCTATTAT 180
 QY 181 GCTAACTAACATTTACCCCTGAAGTTAATGATCATCAAGAGATGAGGCTTAAT 240
 DB 181 GCTAACTAACATTTACCCCTGAAGTTAATGATCATCAAGAGATGAGGCTTAAT 240
 QY 241 GAATCGCTTATTTGAATTAACAGGTGGATCGTTCTGTTGATGATCTTCTTGG 300
 DB 241 GAATCGCTTATTTGAATTAACAGGTGGATCGTTCTGTTGATGATCTTCTTGG 300
 QY 301 CGTGGGAGTCACATTTGACACATATCAGCCACTGAAACAGAGTCTGGTACAAACAA 360
 DB 301 CGTGGGAGTCACATTTGACACATATCAGCCACTGAAACAGAGTCTGGTACAAACAA 360
 QY 361 GTAAGCTGTCACATTTGAGAGCATATGTCGCCGCGGTCATGTTTCACTTGTG 420
 DB 361 GTAAGCTGTCACATTTGAGAGCATATGTCGCCGCGGTCATGTTTCACTTGTG 420
 QY 421 GATATTAACACGAGAAATTTATATCTTGTTCATGTTGTTGATTTGACGATAGTA 480
 DB 421 GATATTAACACGAGAAATTTATATCTTGTTCATGTTGTTGATTTGACGATAGTA 480
 QY 481 A 481
 DB 481 A 481

RESULT 2
 AE000258 13172 bp DNA linear BCF 01-DEC-2000
 LOCUS Escherichia coli K12 MG1655 section 148 of 400 of the complete genome.
 DEFINITION
 ACCESSION AE000258 U00096
 VERSION AE000258.1 GI:2367121
 SOURCE
 ORGANISM
 Escherichia coli K12.
 Escherichia coli K12.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 1 (bases 1 to 13172)
 Blatter, F.R., Plunkett, G., III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Colado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., and Shao, Y.

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 9278503
 2 (bases 1 to 13172)
 Blatter, F.R.
 Direct Submission
 Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459
 3 (bases 1 to 13172)
 Blatter, F.R.
 Direct Submission
 Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459
 4 (bases 1 to 13172)
 Plunkett, G. III.
 Direct Submission
 Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 On Sep 9, 1997 this sequence version replaced gi:1787907.
 This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: markborov@ga.tech.edu]. Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (http://cgsc.biology.yale.edu). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

FEATURES
 source
 gene
 CDS
 Location/Qualifiers
 1. 13172
 /organism="Escherichia coli K12"
 /strain="K12"
 /sub_strain="MG1655"
 /db_xref="taxon:83333"
 76..1668
 /gene="malX"
 /note="b1621"
 76..1668
 /gene="malX"
 /function="Transport; Transport of small molecules: carbohydrates, organic acids, alcohols"
 /note="o530; 99 pct identical to PROA-ECOLI SW: P19642; CG Site No. 32254; phosphotransferase system enzyme II"
 /codon_start=1
 /transl_table=11
 /product="PTS system, maltose and glucose-specific II ABC"
 /protein_id="AAC74693.1"
 /db_xref="GI:1787908"
 /translation="MTAKAPKVTLMEEFPOOLGKFTMLPVALLSFCGIMLGSSLS HVITLIPVLGNPVIQALFTWMSKIGSFATSLPVMFCIALPLGARLNKGVAAAFGE

|||||

DB 5454 GATTATTAAACCGAATTATTATCTGTGCATGTTGTGTCGATTCGACGATAGTA 5513

QY 481 A 481

DB 5514 A 5514

RESULT 3

D90806 15332 bp DNA linear BCT 29-MAY-1997

LOCUS E.coli genomic DNA, Kohara clone #315(36.6-36.9 min.).

DEFINITION D90806 AB001340

ACCESSION D90806.1 GI:1742679

VERSION D90806.1 GI:1742679

KEYWORDS Complete and shotgun sequencing; add: hsdh, hsdh; malI; malY; malY; ntl.

SOURCE Escherichia coli (strain:K12) DNA, clone_lib:Kohara lambda miniset library clone:Kohara clone #315.

ORGANISM Escherichia coli

REFERENCE Escherichia coli

AUTHORS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

1 (sites)

Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T., Salto,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y., and Yano,M.

The systematic sequencing of the Escherichia coli genome in Japan Unpublished

2 (bases 1 to 15332)

Mori,H

Direct Submission

Submitted (14-DEC-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5 Takayama, Ikoma, Nara 630-01, Japan

(E-mail:hmori@gtc.aist-nara.ac.jp, Tel:81-7437-2-5660, Fax:81-7437-2-5669)

Collaboration Information:

Project:

The Japan E.coli genome DNA sequencing project

Group:

The Japan E.coli genome DNA sequencing group

Members: (1995.4 - 1996.3)

Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T., Salto,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y., and Yano,M.

Headed by:

Name: Takashi Horiiuchi

Address: National Institute of Basic Biology, Okazaki, 444, Japan

E-mail: kishori@nibb.ac.jp

Information operator:

Name: Hirotsada Mori

Address: NARA Institute of Science and Technology, Ikoma, 630-01, Japan

E-mail: hmori@gtc.aist-nara.ac.jp

URL:

The Japan E. coli genome database

http:bsw3.aist-nara.ac.jp

Location/Qualifiers

1. .15332

/organism="Escherichia coli"

/strain="K12"

/db_xref="taxon:562"

/map="36.6 min"

/clone="Kohara clone #315"

/clone_lib="Kohara lambda miniset library"

CDS

/note="Nucleotide position 1699836-1715167 from the initiation site of Thra (0 min.). This clone is from Kohara lambda miniset library"

complement(<1.48)

/note="ORF_ID:0314#4

similar to [Swissprot Accession Number P34000]"

/codon_start=1

/transl_table=1

/product="Potential acryA operon repressor."

/protein_id="BA15376.1"

/db_xref="GI:1742680"

/translation="TMDNQTEAQTPTRTI"

complement(269..1036)

/gene="hdhA, hsdh"

complement(269..1036)

/gene="hdhA, hsdh"

/note="ORF_ID:0314#5

similar to [Swissprot Accession Number P25529]"

/codon_start=1

/transl_table=1

/product="7-a-hydroxysteroid dehydrogenase (EC 1.1.1.159)

(7.alpha-HSDH)"

/protein_id="BA15377.1"

/db_xref="GI:1742681"

/translation="MFNSDNLRLDGCALITGACAGIGKEIATFATGASVYVSDIN

ADANHYVDEIQQLGQAFACRCDITSEOEISALADPAISLGLVDILVNNAGGGPK

PEDMPMADFERAYELNVESEFHLISQIVAPENKGGVILITISMAENKNNMTSYA

SSKKAASHLVNAFMFCEDGKNIYVNGIAPGAILDLDAKSVITPEIQMKLHPHRL

COPDINAAAFMFCSPASWSGQILVSGGVEQLN"

complement(1148..2176)

/gene="malI"

complement(1148..2176)

/note="ORF_ID:0314#6

similar to [Swissprot Accession Number P18811]"

/codon_start=1

/transl_table=1

/product="Maltose regulon regulatory protein MalI."

/protein_id="BA15378.1"

/db_xref="GI:1742682"

/translation="MTAKKITIDHVALAGVSIVSLVLSGKRISTATGEVNNAA

IEELCFVNNROASALRGQSGVIGLIYRDISAPFYAELTGLTALRAGSRVLLDG

KGDEQLAORFSLNLQGVDEIVLGAAGSSDDLRRMAEKAIPTIRASRYLDVD

TVRPDPMQAOLITHEILIRNGHRIAWLGGSSSLTRAERVGTCATLLKGLPFPHD

WVLECTSSQKQAAEAIITALLRHNPISAVVQNETIAGNFGILLKGRSGESGVIR

YEEQVSLAATFATPTLDDIPVTWASTPAREIGITLADRMOKITHEETHSRNLL

PARLIAAK"

2351..3943

/gene="malX"

2351..3943

/gene="malX"

/note="ORF_ID:0314#7

similar to [PIR Accession Number B42477]"

/codon_start=1

/transl_table=1

/product="Phosphotransferase system enzyme II (EC 2.7.1.69) MalX"

/protein_id="BA15379.1"

/db_xref="GI:1742683"

/translation="MTAKTAPVPTLMEFFOOLGTFMLPVALLSFCGIMLGSSLS

HDVITLPIVGNPVOAIFETWMSKIGSAFSLPMPFCAIPLGLARENKGAFAFG

ICYAVMNIAVNNFWLNNKGIITPTTDAVAVKANNIOSIGIOSIDGILGAVAGIIVNM

LHERHNIRLPDPALEFGGTRFVPIISSLVWGLVGLVPLVPIFAMKISGLKHNIS

AGDFGPMLEGTGERLLLPFGLHHIIVALLIRFTDAGTQEVCGQVSGALTFQALDSC

PTTGFSESATREFLSQGMKPAFELGGLPGALAMYHCAPENRHKIKGLISGLLACV

GGTEPLELEFLFVAPVLYVYIHALLTGGLFTVMSVGLVITGNTDGNITIDFVFGILHG

ISTKRWMPVVAIWEVYVYIFRAITRPNLKTGGRDSEVASTEKAVACAPGSGY

NYPATILEAGADNTVSDNCTIRLRISVKKMSLVNQAALDNRAIGVQVQNGHMLVY

VIGPQVQSVKDMAGLMTVQA"

3953..5125

/gene="malY"

3953..5125

/gene="malY"

/gene="malY"

CDS

```

/note="ORF_ID:0314#8
similar to [PIR Accession Number C42477]"
/codon_start=1
/transl_table=1
/product="Walt protein"
/protein_id="BA15380.1"
/db_xref="GI:1742684"
/translation="MFDESKVVDHRRHGTWCOTMDYVADRFETADLLPTISMDPATAP
CIIFALNORLMHGVGYSRMKNDPEFLIAIAMPSTOHTAIDISQTVYGPSPVITWSE
LIRQMSSEGEVYHTPAYIDAFYKAIBENQNTWPALEKADGDFCDMGKLEAVLAK
PECKIMLCSPQNTGKVTCDDELEIMADLERGVRIISDEIHMDYWGQPHIPLWS
NARGDAMLLTSGSKSEFIPALTGAYGIIENSSRDADLSALKRSDGLSSPSVALTA
HIAAYOQAPWLDALRIYIKDNLTYIADKNAAAPPELMQIPQSTYLAIDLRLNID
DNALOKALIEQEKVAIMPGTYGEGRGFVRLNMGCPRSKLEKGYAGLINRAVR"
5229..6230
/gene="add"
/gene="add"
5229..6230
/gene="add"
/note="ORF_ID:0314#9
similar to [SwissProt Accession Number P22333]"
/codon_start=1
/transl_table=1
/product="Adenosine deaminase (EC 3.5.4.4) (Adenosine
aminohydrolase)."
/protein_id="BA15381.1"
/db_xref="GI:1742685"
/translation="MIDTTLPLDTHRHLDGNIRPORTLEIROYNISLPAOSLETL
PHVOYIANEPDVSFLTKLDMGVYKASLDCRRAVENIDAAHGHLYELRSPG
YMAAHQPLVAGVYEAVIDGVREGKTFGVAKLIGISRFGEACQOLEAFLAHR
DOITALDLAGELGEPGSLSHFNRAADAGMHITVAGSEAGPESIMQALRELAER
IGHGVKALIEDRALMDFLAEQOIGIESCLTISQISVLAELAHPLKTFLEHGRASIN
TDDGVGVGDVRIHEHYVAAPAGLSREQIRQAOINGLEMATLSAEKRALREKVAK"
complement(6264..7343)
/note="ORF_ID:0314#10
similar to [SwissProt Accession Number P37168]"
/codon_start=1
/transl_table=1
/product="Virulence factor Mv1M."
/protein_id="BA15382.1"
/db_xref="GI:1742686"
/translation="MSNSHRTSMKNDSDNIRVGLIGYGYAKTFEHAPLIAGTPGOE
LAVISSDETQVKAQWPTVYVSEPKHLFNDPNIDLIYIPNDTHPELAAALEAGK
HVVYDKPTTVLSQARELDALAKSLGRVLSYFNHRMSDFLLTGLAEVLGEVAY
FESHDRPQVDRMRROGGSGSIWYDLAPHLDDAITLFGLPVSMYVLAOLRPG
AQSTDYFHAIISYQORYILHGTMLAAAESRYIVHSGRSYVKGIDPOERLNGE
RLPOEDMGYDRDVLTRVEGEVEETLLTPGNYPAIYAIRDALNGDENPVPAS
QAIVMELEIGIESAKHRAFLCLA"
9344..9922
/note="ORF_ID:0316#4
similar to [SwissProt Accession Number P00211]"
/codon_start=1
/transl_table=1
/product="Ferredoxin II."
/protein_id="BA15383.1"
/db_xref="GI:1742687"
/translation="MNAWIVAAVVSILGLAFALIGASRPVADDPVYEKIDELI
POSQGGCGTGPCRPYPAISGNGEKNRCRPGGAVALKIAELINVEPOPDEAOE
ITPARVAVIDENNCTGCTKCIACPDVAIVIGAIRAMHTVMSDLCTGNCICVDPPIH
CISIQPAVEIPDPDSKMDINTIPRIIPEVHHA"
9915..12137
/note="ORF_ID:0316#5
similar to [SwissProt Accession Number P14922]"
/codon_start=1
/transl_table=1

```

Query Match 97.5%; Score 469; DB 1; Length 15332;
 Best Local Similarity 99.8%; Pred. No. 2,4e-112;
 Matches 480; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 TGCCTCTCATTTGAGTATGATGCTATTTGACATATCCTTTACCCAGCCTCAAGAGTTT 60
 DB 7310 TGCCTCTCATTTGAGTATGATGCTATTTGACATATCCTTTACCCAGCCTCAAGAGTTT 7369

```

OY 61 AATAACCTCCAGCATATAAGGATGTGTTTAACTTAAGTCAAAAAAATAGCGAATTTTC 120
DB 7370 AATAACCTCCAGCATATAAGGATGTGTTTAACTTAAGTCAAAAAAATAGCGAATTTTC 7429
OY 121 CAACGACAAAAGCTAAATATATGCAAAAAACCTCAGTAAATCTTGCTGAGACTATTAT 180
DB 7430 CAACG-ACAAAAGCTAAATATATGCAAAAAACCTCAGTAAATCTTGCTGAGACTATTAT 7488
OY 181 GCTAAGTAACTTTACCCCTCAAGTAAATGATCAATCAAGAGATGTGGCTGTAA 240
DB 7489 GCTAAGTAACTTTACCCCTCAAGTAAATGATCAATCAAGAGATGTGGCTGTAA 7548
OY 241 GAATGCTCTTATGATTAATTAACAGGTGTGATGCTGTGCTGCTTGAAGTATCTTCTTG 300
DB 7549 GAATGCTCTTATGATTAATTAACAGGTGTGATGCTGTGCTGCTTGAAGTATCTTCTTG 7608
OY 301 CGTGGCGAGTCACATTTGACACATATACAGCCACCTGACAGAGTGTCTGCTGACACACA 360
DB 7609 CGTGGCGAGTCACATTTGACACATATACAGCCACCTGACAGAGTGTCTGCTGACACACA 7668
OY 361 GTAAGCTCTGCACTTGTGGAGCGACATGCTGCCGCTCGGCTGATGTTTCACTTGTG 420
DB 7669 GTAAGCTCTGCACTTGTGGAGCGACATGCTGCCGCTCGGCTGATGTTTCACTTGTG 7728
OY 421 GATATTAACCGAGATTTATATATCTGTTGATGTTGTTGATGTTGACGAGATGTA 480
DB 7729 GATATTAACCGAGATTTATATATCTGTTGATGTTGTTGATGTTGACGAGATGTA 7788
OY 481 A 481
DB 7789 A 7789

```

RESULT 4
 D90808 15601 bp DNA linear BCT 29-MAY-1997
 LOCUS E.coli genomic DNA, Kohara clone #317(36.6-36.9 mln.).
 DEFINITION D90808 AB001340
 ACCESSION D90808.1 GI:1742711
 VERSION Complete and shotgun sequencing; add: hbha, hsdh; mal; malx; maly;
 KEYWORDS Complete and shotgun sequencing; add: hbha, hsdh; mal; malx; maly;
 ntl.

SOURCE Escherichia coli (strain:K12) DNA, clone_1lb:Kohara lambda miniset
 library clone:Kohara clone #317.

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia coli

REFERENCE 1 (sites) Alba,H., Baba,T., Fujita,K., Hayashi,K., Inada,T., Isono,K.,
 Itoh,T., Kasai,H., Kashimoto,K., Kimura,S., Kitakawa,M.,
 Kitagawa,M., Makino,K., Miki,T., Mizobuchi,K., Mori,H., Mori,T.,
 Motomura,K., Nakade,S., Nakamura,Y., Nashimoto,H., Nishio,Y.,
 Oshima,T., Saito,N., Sempel,G., Seki,Y., Sivasubram,S.,
 Tagami,H., Takeda,J., Takemoto,K., Takeuchi,Y., Wada,C.,
 Yamamoto,Y. and Horiiuchi,T.
 A 570-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28,040,1 min region on the linkage map
 DNA Res. 3 (6), 363-377 (1996)

TITLE JOURNAL MEDLINE REFERENCE
 97251357

AUTHORS Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiiuchi,T.,
 Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K.,
 Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M.,
 Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H.,
 Motomura,K., Nakamura,Y., Nashimoto,H., Nishio,Y., Oshima,T.,
 Saito,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,
 Yamamoto,Y. and Yano,M.

TITLE JOURNAL REFERENCE
 3 (bases 1 to 15601)

TITLE JOURNAL REFERENCE
 3 (bases 1 to 15601)

TITLE JOURNAL REFERENCE
 3 (bases 1 to 15601)

TITLE JOURNAL REFERENCE
 3 (bases 1 to 15601)

TITLE JOURNAL REFERENCE
 3 (bases 1 to 15601)

TITLE JOURNAL REFERENCE
 3 (bases 1 to 15601)

TITLE JOURNAL REFERENCE
 3 (bases 1 to 15601)

COMMENT

Takayama, Ikoma, Nara 630-01, Japan
(E-mail:hmori@tc.aist-nara.ac.jp, Tel:81-7437-2-5660,
Fax:81-7437-2-5669)
Collaboration Information:

Project:

The Japan E.coli genome DNA sequencing project

Group:

The Japan E.coli genome DNA sequencing group

Members: (1995.4 - 1996.3)

Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A.,
Horuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S.,
Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S.,
Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K.,
Masuda,S., Miki,T., Mizoduchi,K., Mori,H., Motomura,K.,
Nakamura,Y., Nashimoto,H., Nishio,Y., Oshima,T., Saito,N.,
Sampel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,
Yamamoto,Y. and Yano,M.

Headed by:

Name: Takashi Horuchi

Address: National Institute of Basic Biology, Okazaki, 444, Japan

E-mail: kishori@nibb.ac.jp

Information operator:

Name: Hirotsada Mori

Address: NARA Institute of Science and Technology,

Ikoma, 630-01, Japan

E-mail: hmori@tc.aist-nara.ac.jp

URL:

The Japan E. coli genome database

http:bsw3.aist-nara.ac.jp.

Location/Qualifiers

1. 15601

/organism="Escherichia coli"

/strain="K12"

/db_xref="taxon:562"

/map="36.6 min"

/clone="Kohara clone #317"

/clone_id="Kohara lambda miniset library"

/note="Nucleotide position 1699735-1715335 from the

initiation site of Thra (0 min.).-This clone is from

Kohara lambda miniset library"

complement(1. 149)

/note="ORF_ID:0314#4

similar to [SwissProt Accession Number P34000]"

/codon_start=1

/transl_table=11

/product="potential acraB operon repressor."

/protein_id="BA15406.1"

/db_xref="GI:1742712"

/translation="MMDMOTEAQPTRTILNAREIFSENGFHSASKAKSCAIS

PGTLY"

complement(370. 1137)

/gene="hdbA, hsdH"

complement(370. 1137)

/gene="hdbA, hsdH"

/note="ORF_ID:0314#5

similar to [SwissProt Accession Number P25529]"

/codon_start=1

/transl_table=11

/product="7-a-hydroxysteroid dehydrogenase (EC 1.1.1.159)

/protein_id="BA15407.1"

/db_xref="GI:1742713"

/translation="MFNSDNLRLDGKCAITAGAGIGKEIAITFANAGASVVDIN

ADANHVYDEIQQLGQAFACRCDITSQELSLADLAFISLKVYDILVNNAGGGPK

PEDMADFERAYENFVSFPHLSQIVPEKNGGVILITTSMAENKINNTSYA

SSKAASHLVNNMAFDIGKRNITNGTAPGAILTDALKSVTITPECKMLQHTIRRL

GQPDIAAAALFLCSPPASVNSGQILTVSGGVDPLN"

complement(1249. 2277)

/gene="mali"

complement(1249. 2277)

/gene="mali"

complement(1249. 2277)

/note="ORF_ID:0314#6

similar to [SwissProt Accession Number P18811]"

/codon_start=1

/transl_table=11

/product="Maltose regulon regulatory protein Mali."

/protein_id="BA15408.1"

/db_xref="GI:1742714"

/translation="MAFKKITIHVLAAGVSVSVLSLGSGRISRTGGEVNA

IEELGIVRNROASALRGOSGVIGLITRDLSPFVPEALFAGLEALAEAGRWFLHG

GKDRQALRFSILLNOGVDAVIAAGSSDDLRAEAEKAIPVIFASRASYLDVD

TVRPDMQAQILTEHLIRNGHRIALMGOSSSLRAERNGCATLLKFLPFSHD

VWECSVSQAQAEATLALRNPTISAVACINETIAMGAMGLKAGOSSEGYDR

VEQOVSLSAFDPTTLDIPVWASIPARELITTLADRMQKITHEHSRLNI

PARLIAX"

2452. 4044

/gene="maliX"

2452. 4044

/gene="maliX"

/note="ORF_ID:0314#7

similar to [PIR Accession Number B42477]"

/codon_start=1

/transl_table=11

/product="phosphotransferase system enzyme II (EC

2.7.1.69) MaliX"

/protein_id="BA15409.1"

/db_xref="GI:1742715"

/translation="MTAKTAPKVTTLMEFQQLGKTFMPLVALSPCCIMIGISLSS

HDVITLIPVGNFVLOAIFTWMSKIGSFAPFVMECAIPGLARENKGYAAEAGF

IGYAVNNIAVNFNLTKGILPTDAVLANNNISILGOSIDTGLGAVIGIITVM

IERPHNIRLPDALARFGGTRFVPIISLVMGVGLVTLMPIFAMGISGAMINS

AGDGPMLBGTGRLRLPRGLHIIVALLIFPDAGSTGCGCTVSGALITGIONOSC

PTTHGSESATRFLSQKMAFLGLGLGALAMTICARBNRRIKIGLISILAIAY

GGTEPLELEFLFVAPVLYVTHALGLGTVAIVGLTIGNTGNIIDVFGKILG

LSTRKMWEPVVAWIMFVYVYVIFRFAITRNLKTPGRDSEVASSIERAAAGAPKGY

NVPAILEALGADNIVSLDNCITRLRLSVKMSLVVWVALKDMRAIGVVOLOHNLQV

VIGPOVSVKDEAGLMHTVQA"

4054. 5226

/gene="maliY"

4054. 5226

/gene="maliY"

/note="ORF_ID:0314#8

similar to [PIR Accession Number C42477]"

/codon_start=1

/transl_table=11

/product="MaliY protein"

/protein_id="BA15410.1"

/db_xref="GI:1742716"

/translation="MFDESKVVRHGTQCTQWYVADREGTADLLEPFTISMDFAFAP

CIEALNORLMHGVSFRMKNDLFLAIAHMFSTOYAIADQTVVYGSVLYVASE

LIRWSEFGVVIHMPAYAPAFKAIEGNQRTVPALEKROAGWFCDMGKLEAVLAK

PECKIMLCSPQNTGKVTCDLELEIMADICRHGVRVTSDEIHMVMVGEOPHIPS

NVARGDALITSGSKFNITALTGATGIIENSSRDPAIYSAIKRGGLSSPSYLAITA

HIAVYQGAFLWDLALRLYLKDNLTYLADKNNALPELNMQIPQSTYLAWLDPNLID

5330. 6331

/gene="add"

5330. 6331

/gene="add"

/note="ORF_ID:0314#9

similar to [SwissProt Accession Number P22333]"

/codon_start=1

/transl_table=11

/product="Adenosine deaminase (EC 3.5.4.4) (Adenosine

aminohydrolase) "

/protein_id="BA15411.1"

/db_xref="GI:1742717"

/translation="MIDTLPPLDHRHLDGNIRPQITELGRQYNISLPAQSLLETLI

PHYVQINDEDLVSEFLKLMGVYVLAIDLACRRVAFENEDAAHGLHVEVLFSPG

YMAHOLDPAGVVEAVITDVREGCRFFGVQALIGMSTRFEGAGCOOQLAEFLHR

DOITADLADDELGFPSSFLSHNRARDGMHTIYAAELAAPIKFTFDEHGIRASIN

IGHGVKAIEDALMDFTAEQIGISCLTISNIGTISTYAEELAAPIKFTFDEHGIRASIN

ITDPGVGVGDIITHRYTVAAPAGLSRQIRQAOINGLEMAFLSAEERLRERVAAK"

complement(6365. 7444)

/note="ORF_ID:0314#10

similar to [SwissProt Accession Number P37168]"

CDS

gene

CDS

gene

CDS

gene

CDS

gene
CDS


```

/codon_start=1
/transl_table=11
/product="Virulence factor wv1m."
/protein_id="BA15412.1"
/db_xref="GI:1742718"
/translation="MSNSHHTSMRGNDMSDNIRVGLIGYVASKTFHAPLIAGTPEOE
LAVISSDETKVADWPVTYVSEPPHEDNDIVIPFNDFHPLAKALEGK
HYVVKPPTVYLSOARERDIALKSLGRVTSFHNWSDFTLNLGLAEGLGKAY
FESHDFRPPQVDRKRRREGSGSGLWDLAHLDDQATLLEGLPVSMTVDLAQLKPG
AOSTDFHAILSPQRRVILHGTMLAAASARYIVHSGSYVKYGLDPOERLNKGE
RLPOEDWGYMDRGVLTIVGEGERVEETLLTVPGNPAYAYAIRDALNDGENPVPAS
QAIVMELIELIESAKHRATLCIA"
9445.10023
/contig="ORF_ID:0316#4
similar to [SwissProt Accession Number P00211]"
/codon_start=1

Query Match      97.5%; Score 469; DB 1; Length 15601;
Best Local Similarity 99.8%; Pred. No. 2.4e-112;
Matches 480; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 TGCCTCATTTGAATGATGCGCTATTGACACTATCCTTACCCAGCTCAACAGTTT 60
    |||||||
DB 7411 TGCCTCATTTGAATGATGCGCTATTGACACTATCCTTACCCAGCTCAACAGTTT 7470

QY 61 AATACCTGCGACCAATAGGAGATGTTGTTAACTTAAGTCAAGCAAAATAGCGAATTTTC 120
    |||||||
DB 7471 AATACCTGCGACCAATAGGAGATGTTGTTAACTTAAGTCAAGCAAAATAGCGAATTTTC 7530

QY 121 CAACGAACAAGCTAAATATCGCAAAACCTGAGTAAATATCTGCTGAGCTATTATT 180
    |||||
DB 7531 CAACG-ACAAAGCTAAATATCGCAAAACCTGAGTAAATATCTGCTGAGCTATTATT 7589

QY 181 GCTAAGTAACTTATACCCCTGAGATTAATGATCAATCAAGAGATGTGGCTGTAAAT 240
    |||||
DB 7590 GCTAAGTAACTTATACCCCTGAGATTAATGATCAATCAAGAGATGTGGCTGTAAAT 7649

QY 241 GAATCGCTTATGAAATTAACAGTGTGATGCTTCTGTGTTCACTATTTCTTCTG 300
    |||||||
DB 7650 GAATCGCTTATGAAATTAACAGTGTGATGCTTCTGTGTTCACTATTTCTTCTG 7709

QY 301 CGTGGGAGTACATTTGACACTATGACGACCTGAACAGATGCTTGGTACACACAA 360
    |||||||
DB 7710 CGTGGGAGTACATTTGACACTATGACGACCTGAACAGATGCTTGGTACACACAA 7769

QY 361 GTAAGCTCTGCACTGTGAGCAGCATGCTGCCGTCGGGTCATGTTTCACTTGTG 420
    |||||||
DB 7770 GTAAGCTCTGCACTGTGAGCAGCATGCTGCCGTCGGGTCATGTTTCACTTGTG 7829

QY 421 GATATTAAACAGCAATTTTATCTTGTGATGTTGTGATGTTGTGAGGATTAATA 480
    |||||||
DB 7830 GATATTAAACAGCAATTTTATCTTGTGATGTTGTGATGTTGTGAGGATTAATA 7889

QY 481 A 481
    |
DB 7890 A 7890

RESULT 5
LOCUS      D90807              20231 bp      DNA      linear      BCT 29-MAY-1997
DEFINITION E.coli genomic DNA, Kohara clone #316(36.7-37.1 min.).
ACCESSION  D90807 AB001340
VERSION     D90807.1 GI:1742692
KEYWORDS   Complete and shotgun sequencing; add; gsc; nth; pdxH; lys; ydha;
            yhcp.
SOURCE      Escherichia coli (strain:K12) DNA, clone_11b:Kohara lambda miniset
            library clone:Kohara clone #316.
ORGANISM    Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1 (sites)
AUTHORS     Aliba,H., Baba,T., Fujita,K., Hayashi,K., Inada,T., Isono,K.,
            Itoh,T., Kasai,H., Kashimoto,K., Kimura,S., Kitakawa,M.,
            Kitagawa,M., Makino,K., Mori,H., Mizobuchi,K., Morita,H.,
            Motomura,K., Nakade,S., Nakamura,Y., Nishimoto,H., Nishio,Y.,
            Oshima,T., Saito,N., Sampel,G., Seki,Y., Sivasubaram,S.,
            Tagami,H., Takeda,J., Takemoto,K., Takeuchi,Y., Wada,C.,
            Yamamoto,Y. and Horiuchi,T.
            A 570-kb DNA sequence of the Escherichia coli K-12 genome
            corresponding to the 28.0-40.1 min region on the linkage map
            DNA Res. 3 (6), 363-377 (1996)
            97251357
TITLE       2 (sites)
JOURNAL     2 (sites)
MEDLINE     2 (sites)
REFERENCE   2 (sites)
AUTHORS     Aliba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiuchi,T.,
            Ikehata,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K.,
            Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M.,
            Kitakawa,M., Makino,K., Masuda,S., Mori,H., Mizobuchi,K., Morita,H.,
            Motomura,K., Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T.,
            Saito,N., Sampel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,
            Yamamoto,Y. and Yano,M.
            The systematic sequencing of the Escherichia coli genome in Japan
            Unpublished
            3 (bases 1 to 20231)
            Mori,H.
            Direct Submission
            Submitted (14-DEC-1996) Hirota Mori, NARA Institute of Science
            and Technology, Res. & Edu. Center for Genetic Info.; 8916-5
            Takayama, Ikoma, Nara 630-01, Japan
            (E-mail:hmorigtc.aist-nara.ac.jp, Tel:81-7437-2-5660,
            Fax:81-7437-2-5669)
COMMENT     Collaboration Information:
            Project:
            The Japan E.coli genome DNA sequencing project
            Group:
            The Japan E.coli genome DNA sequencing group
            Members: (1995.4 - 1996.3)
            Aliba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A.,
            Horiuchi,T., Ikehata,K., Inada,T., Isono,K., Isono,S.,
            Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S.,
            Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K.,
            Masuda,S., Mori,H., Mizobuchi,K., Morita,H., Motomura,K.,
            Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T., Saito,N.,
            Sampel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,
            Yamamoto,Y. and Yano,M.
            Headed by:
            Name: Takashi Horiuchi
            Address: National Institute of Basic Biology, Okazaki, 444, Japan
            E-mail: kishori@nibb.ac.jp
            Information operator:
            Name: Hirota Mori
            Address: NARA Institute of Science and Technology,
            Ikoma, 630-01, Japan
            E-mail: hmorigtc.aist-nara.ac.jp
            URL:
            The Japan E. coli genome database
            http:bsw3.aist-nara.ac.jp.
            Location/Qualifiers
                1..20231
                    /organism="Escherichia coli"
                    /strain="K12"
                    /db_xref="taxon:562"
                    /map="36.7 min"
                    /clone="Kohara clone #316"
                    /clone_11b="Kohara lambda miniset library"
                    /note="Nucleotide position 1705226-1725456 from the
                    initiation site of Theta (0 min.). This clone is from
                    Kohara lambda miniset library"
                    1..840
                        /gene="add"
                        <1..840
                            /gene="add"
                            /note="ORF_ID:0314#9
                            similar to [SwissProt Accession Number P22333]"
                            /codon_start=1
                            /transl_table=11

```

/product="Adenosine deaminase (EC 3.5.4.4) (Adenosine
aminohydrolase)."
/protein_id="BA15388.1"
/db_xref="GI:1742693"
/translation="DIVSELFLKIDMGVVLASLDACRRVAFENIEDAARHGLHYELR
FSPGMAHOLPVAQVENVIDCVRECCGPGVQAKTIGMSRTFGGAACQOLEAF
LAHQDTALDLAGEDEFGPSLSLSHRNRARDAQMTHTVAGACAGESTWQAIREL
GAERIGGVKAIEDRALDLAEQOIGESCLTSIQSTVAELAHPLKTFLEHGIR
ASINTDPGVGVVDIHEHYTVAAPAGLSRBOIRQAOINGLEMAFLSAEKRALREKY
AAK"
complement(874..1953)
/note="ORF_ID:0314#10
similar to [SwissProt Accession Number P37168]"
/codon_start=1
/transl_table=11
/product="Viralence factor MYM."
/protein_id="BA15389.1"
/db_xref="GI:1742694"
/translation="MSNSHTSMRGNDSNDNRVGLIGYVAKTEFHAFLIAGPGE
LAVISSDETKADMPVTYVYSEPKHIFNDPNIDLIYIPPNDRHPLAKALAEAK
HVVYDKPPTVLSQARELDALAKSLGKLVFHNKRWSDLETLGLAEGLGVAY
FESHDRFRPOVRWRBQGGSGIWDLAHLLDQAITLGLFVSWTVDLAQLRG
AOSTDYFALLISYPOARVILHGTMTAAESARYIVHSGRSYVKKGLDPSDELAE
RLPQEDMGDMRBDGLTVREGEVEETLLVPGNYPAYVAIRDALNGDENPVPAS
OAIQVMELEIGIESAKHRATLCLIA"
3954..4532
/note="ORF_ID:0316#4
similar to [SwissProt Accession Number P00211]"
/codon_start=1
/transl_table=11
/product="Ferredoxin II."
/protein_id="BA15390.1"
/db_xref="GI:1742695"
/translation="MAWIAVAANVSLGLAALGASRRFVVEDPVRVYKIDELL
POSQCGQCGPCGRPAEAIISGNERINCCAGGAEVMLKIAELLNVEPQIDGAE
ITPARMAVITDENNCIGCTKCIQACPVDAIVGATRAMHTVMSDLCTGCMLCVDPPTH
CISIQVAETPDPSMKMDLNTIPRILIVPEHHA"
4525..6747
/note="ORF_ID:0316#5
similar to [SwissProt Accession Number P14922]"
/codon_start=1
/transl_table=11
/product="Glucose repression mediator protein."
/protein_id="BA15391.1"
/db_xref="GI:1742696"
/translation="MKLFSAFRKNIKIMPNGGIIHPEMKQSNGLPRLVPLARV
ITPKHIGAGEGLCYVGRKVLRGQPLTGRGKMLPVHPTSGTYATALPSTAHSA
LAELSVLIDADEDCWIPRDGWDYRTSRRELEIRIHOFVAGLGGAGFPPTGKLG
GGDKLETLIINAEECPYITADRLMDCAQVVEIRILAHILDRRETLIGIDKRP
OAIISMLRAVLADSNDSLIVITPKYPSGAKQULYITLKQVPHGGRSSDIEVLQNV
GTAAYAKRAVIDGEPIITERVTLTGEAIAIRPGVMAIRLCTPVHLLINDAGFCPSADOM
VIMGPGIMGFTLPMVLDVYVKITNCLLASANVLEIGRPOEGSIRCSACADCPADL
POOLYWFSGQCHDKRTTNINADICCEGCANVCPSPNIPINOFREKAEIARQEE
KRAAEKARFEARQARLEREKARLERHRSAAVOPAKKDKHIAALAVYKQKQAT
OPTIVKAGERPNNSAIIARERAKQAQARQAELOQTNDPAANVDEEDVDEKALARA
KARLEQOQANAEPEQVDPKRAAVEAAIARAKARLEQOQANAEDEEDVDEKARA
AATARAKARLEQOQANAEPEQVDPKRAAVEAAIARAKARREQOPANAEEQVDP
RKAAVEAATAARAKARLEQOQANAEPEQVDPKRAAVEAAIARAKAKKAAQKXVVED
"
7810..8430
/note="ORF_ID:0316#10
similar to [SwissProt Accession Number P44291]"
/codon_start=1
/transl_table=11
/protein_id="BA15392.1"
/db_xref="GI:1742697"
/translation="MLKTIKRGITLALPAGSTGLTAAINQMTAKTTTIAEQSLQQA
LFDQVLPABERYNNALASCYLTAPELGEGEHVYIAKODKCAVAVLEATAPDGSYG
AIOLLYAGDPNGVLTCTRVTEHHTPGIDKIELRLSDITTHPAGKKISGADAHNAV
KRDGDFDQFTGATITTPRAVNVAKRAGLTAQTLPAQLSOLPACGE"
8434..9129
/note="ORF_ID:0316#11
similar to [SwissProt Accession Number P43959]"

gene
CDS
/codon_start=1
/transl_table=11
/protein_id="BA15393.1"
/db_xref="GI:1742698"
/translation="MSEIKDVIYQGIKRNNSALVOLGLICPLIAVTSTATNALGLIA
TTVALVTNLTITSLRHTPAEIRIPYVMIASVSAVOMLINAFAGLYOSLIGLEI
PLVTNCTIVVGRAEFAAKGAPALSDAGDSIGMCATCAFEVGSIRELIGNSTLDG
ADALIGSMARVIRELPHDPSFLMLLPFGAIFGLIGLMAGRYLIDERRKKRRAEVA
AERLAPNGEIGNV"
9129..9764
/gene="nth"
9129..9764
/note="nth"
9129..9764
/note="ORF_ID:0316#12
similar to [PIR Accession Number A32412]"
/codon_start=1
/transl_table=11
/product="Deoxyribonuclease (pyrimidine dimer) (EC
3.1.25.1)"
/protein_id="BA15394.1"
/db_xref="GI:1742699"
/translation="MNRKARLEILTRLRENNPHPTLENFSSPELLIAVLSNOAT
VSNKATANKLEPVANTRPAKMLECEGVKTYIKTIGXNSKAKNITCTGCTIILEONG
EVPDRALALELPVGRKTAENVLNTAFGFTLVADTHIRVCRRTOFPAGKVEQVE
EKLKVPFAEFKVDCHHMLLHGHTCIARRKRCGSCITIEDLEYKEKVDI"
10375..11877
/note="ORF_ID:0316#13
similar to [SwissProt Accession Number P36574]"
/codon_start=1
/transl_table=11
/product="Dl-crIpeptide transporter."
/protein_id="BA15395.1"
/db_xref="GI:1742700"
/translation="MSTANOKRTESSVSLNAPKOPAKYLFISIELMERFGYIGLOGIM
AVLYIKOLGSEADSIITFSSFSALVYGLVIAIGMDPKVYGTGRVLMGLAIVLAIGY
ALVANSMDAETVYGNMAIIVNGSLKBRANSSILSTCYKENDRLDGAETMYMSVY
IGSFSSMATPWLAKTGWVSFAFSLVGLITTYNTRFCORRYKQYGRDPDEPIYX
RNLLTLITIGVALLAIATWLLHNOEVAMALGVAFGLIVTEGKEATRKMGARKKAI
VAFILMLEAIIFEVLYSOMPTSLNFEFAIRNVEHLIGLAFASPEQYOLNPEWIIIGSP
ILAIYNNKMGDTLPMPRTKFAIGMYMCSGAPLILPLKAFASDAGIVSVMVASYGLO
SIGELMISGLGLAWAQLVQPRLMGFIINGSFELTAGNLIGYVAGMMAVPDWTDP
"
Query Match 97.5%; Score 469; DB 1; Length 20231;
Best local Similarity 99.8%; Pred. No. 2,4e-112;
Matches 480; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Db 1 TGGCTCTCAATTGAAGTATGATGGCTTTGACACTATCTTTACCAAGCTCAAGAGTTT 60
1920 TGGCTCTCAATTGAAGTATGATGGCTTTGACACTATCTTTACCAAGCTCAAGAGTTT 1979
61 AATAACCTGCAGCAATTAAGGATGTTGTTTAACCTTAAGTCAAAAAAATAGCAAAATTTTC 120
1980 AATAACCTGCAGCAATTAAGGATGTTGTTTAACCTTAAGTCAAAAAAATAGCAAAATTTTC 2039
121 CAACGAACAAGCTAAATATGCCAAAACCTCAGTAAAAATCTGCTGAGCTATTAT 180
2040 CAACG-ACAAACCTAAATATGCAAAAACCTCAGTAAAAATCTGCTGAGCTATTAT 2098
181 GCTAAGTACATTACCCCTGAAAGTTAATGATCAATCAAGAGAGATGGGCTGTAAT 240
2099 GCTAAGTACATTACCCCTGAAAGTTAATGATCAATCAAGAGAGATGGGCTGTAAT 2158
241 GAATCGCTTATTTGAATTAACAGGTTGATCGTTCGTTGCTTTCAGTCAATTCCTCTGG 300
2159 GAATCGCTTATTTGAATTAACAGGTTGATCGTTCGTTGCTTTCAGTCAATTCCTCTGG 2218
301 CGTGGGAGTCACTTGACACTATATAGCCACCTGAACAGAGTGGTTCGTTACACACAA 360
2219 CGTGGGAGTCACTTGACACTATATAGCCACCTGAACAGAGTGGTTCGTTACACACAA 2278
361 GTAAGCTCTGCACTTGGGAGGAGCAATGCGCCGCGCGGAGTCACTTTCACCTGTGCG 420
2279 GTAAGCTCTGCACTTGGGAGGAGCAATGCGCCGCGCGGAGTCACTTTCACCTGTGCG 2338

OY 421 GATATTAAACGAGATTATATCTCTTCGATGTTGTCGATTGTCAGGATAGTA 480
|||||
Db 2339 GATATTAAACGAGATTATATCTCTTCGATGTTGTCGATTGTCAGGATAGTA 2398
OY 481 A 481
|
Db 2399 A 2399

RESULT 6
AE005386
LOCUS
DEFINITION
13259 bp DNA linear BCT 21-MAR-2001
Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 5 of
290
ACCESSION
AE005386 AE005174
VERSION
AE005386.1 GI:12515609
KEYWORDS
SOURCE
ORGANISM
Escherichia coli O157:H7 EDL933.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
AUTHORS
1 (bases 1 to 13259)
Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
Postal,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K.,
Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
Welch,R.A. and Blattner,F.R.
Genome sequence of enterohemorrhagic Escherichia coli O157:H7
JOURNAL
MEDLINE
Nature 409 (6819), 529-533 (2001)
21074935
PUBMED
11206551
TITLE
JOURNAL
Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
Postal,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K.,
Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
Welch,R.A. and Blattner,F.R.
Direct Submission
Submitted (22-OCT-2000) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
JOURNAL
TITLE
Direct Submission
Submitted (22-OCT-2000) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
source
1..13259
location/Qualifiers
/organism="Escherichia coli O157:H7 EDL933"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:155864"
/note="enterohemorrhagic"
118..1710
gene
/gene="malX"
/note="22626"
118..1710
CDS
/gene="malX"
/function="Transport: Transport of small molecules:
Carbohydrates, organic acids, alcohols"
/note="Residues 1 to 530 of 530 are 99.43 pct identical to
residues 1 to 530 of 530 from Escherichia coli K-12 Strain
MG1655: B1621"
/codon_start=1
/transl_table=11
/product="PTS system, maltose and glucose-specific II ABC"
/protein_id="AG5610.1"
/db_xref="GI:12515610"
/translation="MTAKTAPKVTLMFEFOOLGKTFMPLVALLSFCGIMLGSSISS
HDVITLPLVGNPVLOALFTWMSKIGSFAFSLPVMFCIALPLGLAREKGVAAAFGE
GVAYAMNLAVNPNLTKNGILPTDDAVALKANNIOIILICISDITIGAVAGIIVWM
LHERHNIRLDPALAFPGGTRVPVPISSIVMGLVGLVPLVPIRAMGISGIGHINS
AGDGPMLFGTEERLLDPGLHHIIVALLIRFPDAGTQECQIVSGALITFOALSC
PTTHGSSASRIGESQGMKPAFLGLPGALALAMYKARBNRHKIKGLIISGLIACVY
GGTEPLEFLFLFVAPVLYVIALITGGLFTVMVSVIGTGNIDNIIDFVYVGLHG
LSTKWMVPVVVAIWFEVYVYIFRFAITRFNLKTPGRDSEVASSIEKAVAGAPKSGSY

gene
1720..2892
/gene="malY"
/note="22627"
1720..2892
CDS
/gene="malY"
/function="enzyme; Degradation of small molecules: Carbon
compounds"
/note="Residues 1 to 390 of 390 are 99.23 pct identical to
residues 1 to 390 of 390 from Escherichia coli K-12 Strain
MG1655: B1622"
/codon_start=1
/transl_table=11
/product="enzyme that may degrade or block biosynthesis of
endogenous mal inducer, probably aminotransferase"
/protein_id="AG5611.1"
/db_xref="GI:12515611"
/translation="MFDPSKYVDHRGTCQTDQWYADBRGTADLLPTISDMFATP
CITELNORLMHGVGRYRKNDERLAIAMHFTSTOHTATIDPOSVVGPSTIVASE
LIRQWSETEGVVHTTPAYDAFYKAIENQRTVMVPALEKQDGMFCDMGLEAVLAK
PECKIMLCSQNPQKVTWCDELEIMADLERHGVRIYSDIHMNDMVGEOPIHPWS
NVARQDMALITSGSKSFNIPALTGAGTIENESSDAYLSALKGDLSPSLALTA
HIAVQOGAPMIDALRYVTKNLTLYIAQKMNAPPELNMQIPOSTYIAMLDRLPIND
DNALOKATIEQEKAIMRGYTYGEGRGFVRLNACPRSKLEKGVAGLINAIRVR"
2996..3997
gene
/gene="add"
/note="22628"
2996..3997
CDS
/gene="add"
/function="enzyme; Central intermediary metabolism:
salvage of nucleosides and nucleotides"
/note="Residues 1 to 333 of 333 are 99.39 pct identical to
residues 1 to 333 of 333 from Escherichia coli K-12 Strain
MG1655: B1623"
/codon_start=1
/transl_table=11
/product="adenosine deaminase"
/protein_id="AG5612.1"
/db_xref="GI:12515612"
/translation="MIDTTLPLPTDIHRHLDGNIRPOTILELGRQYNISLPAOSLETLI
PHVYIAEPLDVSFLTKLDMGVYIASLADACRYAFENIEDAANGLHYVLEFSPG
YMAAHOLPVGVGVAVIDGVREGCRGTGYOAKIGISRTGEACAOOLEAFAR
DOITALDIADGELGFPGLSFHSFERNRADAQHTTVHNGEAAGPSISQOALELCAER
IGHGVKATIEDRALMDLBOQIGIESCLTSNIQTSVVELAHNPLKTEHETIRASIN
TDDPGVGVSDIIEHYTVAPAGALSREQIQAQINGLEMAFLNAEKKALREKVAK"
complement(4033..5112)
/gene="22629"
complement(4033..5112)
CDS
/gene="22629"
/function="orf; Unknown function"
/note="Residues 1 to 359 of 359 are 99.72 pct identical to
residues 1 to 359 of 359 from Escherichia coli K-12 Strain
MG1655: B1624"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AG5613.1"
/db_xref="GI:12515613"
/translation="MSNSHHTSRGNDMSDNIRVGLIGYAKTFEHPALINGTGLE
LAVISSEDETKKADMPVTVYVSEPKHFNENDINILIVIPPNQDHPPLAKALAGK
HYVDPKPTVITLSQARELDALAKSGRYLSVFNHNRKMSDFITLGLABEVLGEVAY
FESHDFRFPQVDRKRWROGGSGIWTDLAHLDDATTEGLPVSMTVIDLAQLRPG
AOSTDYFHAIISYPOQRVYIRHGTMLAAEASARYIHGSGSYVKGLDPOERLANGE
RLPOEDMGYDMRDGVLTRVEGERVEETLLTPGVNPPAYAIIRDALNGDGNPPVAS
QAIQWMELEIGIESAKRHATFCLA"
5715..5930
gene
/gene="22631"
5715..5930
CDS
/gene="22631"
/function="orf; Unknown function"
/note="Residues 1 to 71 of 71 are 100.00 pct identical to
residues 1 to 71 of 71 from Escherichia coli K-12 Strain

```

MG1655: B1625"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG5614.1"
/db_xref="GI:12515614"
/translation="MTVODYLKFKRKISSLESLEKLYDHINVTLTDDQELINMYRAD
HRRRELVSGRRLFLDGLGVPKSVMHVQ"
/gene="6456"
/gene="2632"
/feature="2632"
/feature="2632"
/function="orf: Unknown function"
/note="Residues 1 to 154 of 154 are 98.70 pct identical to
residues 1 to 154 of 154 from Escherichia coli K-12 strain
MG1655: B1626"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG5615.1"
/db_xref="GI:12515615"
/translation="MRRYSGEPMTTTPQIRGILGPIALMLIALLSTLALLXTA
ALSSPQFTLGAQALTTQILMGVSFTIALMWTTLTAFFKRRRCYFKHTITML
LISVLAVKAFAPFSPEDGIAVRQLFTLATALIVPFKRSSHVKATFVNP"
/gene="7114"
/gene="2633"
/feature="2633"
/function="orf: Unknown function"
/note="Residues 1 to 193 of 193 are 99.48 pct identical to
residues 1 to 193 of 193 from Escherichia coli K-12 strain
MG1655: B1627"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG5616.1"
/db_xref="GI:12515616"
/translation="MTDYLLFVGTVLVNNEVLVKEGLCPMGVSKKLEPMANGLA
TPFWTLASICAMLDITWILPLNLTIRTAFLIVAVVQFTFEMVVRKSPVLYRL
IGIRPLITTCNVIGVALININCHNLQSLALGFSNAVGSLSLVMLVFAAIRBLAV
ADVPAPFRGNALITTAGLMSLAEKFGVGLKL"
/gene="7692"
/gene="2634"
/feature="2634"
/function="orf: Unknown function"
/note="Residues 1 to 192 of 192 are 98.95 pct identical to
residues 1 to 192 of 192 from Escherichia coli K-12 strain
MG1655: B1628"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG5617.1"
/db_xref="GI:12515617"
/translation="MNAIMIAVAAYXLLGLAFLGAILGTASRFAYEDDPVYEKIDEL
FOSQCGCGYPCGRPYAEIAISCNKERKRCACGAGVAMKLAELLNVEPQPLDGAOE
LTPARMVAVIDENNICGCTICACPDVAICPATAMHTVMSDLCTGCNLCVDPPTH
CISLOPVAETPSPKMDLNTIPVIRLIPVEHNA"
/gene="9907"
/gene="2636"
/feature="2636"
/function="orf: Unknown function"
/note="Residues 1 to 740 of 740 are 97.83 pct identical to
residues 1 to 740 of 740 from Escherichia coli K-12 strain
MG1655: B1629"

```

Query Match 94.0%: Score 452.2; DB 1; Length 13259;
Best local Similarity 96.3%; Pred. No. 6,2e-108;
Matches 463; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 TGCCTCTCATTTGAAGTATGATGCTATTTGACACTATCCTTACCACGCTCAACAGTTT 60
|||||

```

Db 5079 TGCCTCTCATTTGAAGTATGATGCTATTTGACACTATCCTTACCACGCTCAACAGTTT 5138
QY 61 AATAACCTGCCAGCAATAAGGATGTTGTTAACTTAAGTCACAAAAATAGCAATTTTC 120
|||||
Db 5139 AATAACCTGCCAGCAATAAGGATGTTGTTAACTTAAGTCACAAAAATAGCAATTTT 5198
QY 121 CAACGACAAAAGCTAAATATGCAAAAACCTCAGTAAAAATCTTGCTGGAGCTATTATT 180
| | | | |
Db 5199 CCAACGACAAAAGCTAAATATGCAAAAACCTCAGTAAAAATCTTGCTGGAGCTATTATT 5258
QY 181 GCTAAGTAACTTTATACCCCTGAAGTTAATGATCAATCAAGAGATGAGGCTTAAT 240
|||||
Db 5259 GCTAAGTAACTTTATACCCCTGAAGTTAATGATCAATCAAGAGATGAGGCTTAAT 5318
QY 241 GAATGCTTTATGATTAACAGGTGATGCTTGTCTGTTCACTATCTTCTTGG 300
|||||
Db 5319 GAATGCTTTATGATTAACAGGTGATGCTTGTCTGTTCACTATCTTCTTGG 5378
QY 301 CGTGCGAGTCACATTTGACAACTATCAGCCACCTGACAGAGTCTTCGGTACACAA 360
| | | | |
Db 5379 CGTGCGAGTCACATTTGACAACTATCAGCCACCTGACAGAGTCTTCGGTACACAA 5438
QY 361 GTAACCTCTGACCTTTGGAGCGAGCATGCTCCGTCGGTGCATGTTTCACTTGCG 420
|||||
Db 5439 GTAGCCCTGCGCTTGAGAGGAGATGCTCCGTCAGGTCATGTTTCACTTGCG 5498
QY 421 GATATTAACCGAGAAATTTATCTTGTGATGTTGTTGATGTCAGGATAGTA 480
|||||
Db 5499 GATATTAACCGAGAAATTTATCTTGTGATGTTGTTGATGTCAGGATAGTA 5558
QY 481 A 481
Db 5559 A 5559

```

RESULT 7
AP002558
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS

Escherichia coli O157:H7
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
1 (sites)
Makino,K., Yokoyama,K., Kubota,Y., Yutsudo,C.H., Kimura,S.,
Kurokawa,K., Ishii,K., Hattori,M., Tatsuno,I., Abe,H., Iida,T.,
Yamamoto,K., Ohnishi,M., Hayashi,T., Yasunaga,T., Honda,T.,
Sasakawa,C. and Shinagawa,H.
Complete nucleotide sequence of the prophage VT2-Sakai carrying the
verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
derived from the Sakai outbreak
Genes Genet. Syst. 74 (5), 227-239 (1999)
20198780
2 (sites)
Ohnishi,M., Murata,T., Nakayama,K., Kuhara,S., Hattori,M.,
Kurokawa,K., Yasunaga,T., Yokoyama,K., Makino,K., Shinagawa,H. and
Hayashi,T.
Comparative analysis of the whole set of rRNA operons between an
enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an
Escherichia coli K-12 strain MG1655
Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
20557356
3 (sites)
Yokoyama,K., Makino,K., Kubota,Y., Matanabe,M., Kimura,S.,
Yutsudo,C.H., Kurokawa,K., Ishii,K., Hattori,M., Abe,H., Iida,T.,
Yamamoto,K., Hayashi,T., Yasunaga,T., Honda,T., Sasakawa,C. and
Shinagawa,H.
Complete nucleotide sequence of the prophage VT1-Sakai carrying the
Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli

0157:H7 strain derived from the Sakai outbreak
Gene 258 (1-2), 127-139 (2000)

JOURNAL
MEDLINE
20364182
4 (sites)
Hayashi,T., Makino,K., Ohnishi,M., Kurokawa,K., Ishii,K.,
Yokoyama,K., Han,C.-G., Ohtsubo,E., Nakayama,K., Murata,T.,
Tanka,M., Tobe,T., Iida,T., Takami,H., Honda,T., Sasakawa,C.,
Ogasawara,N., Yasunaga,T., Kuhara,S., Shiba,T., Hattori,M. and
Shinagawa,H.
Complete genome sequence of enterohemorrhagic Escherichia coli
0157:H7 and genomic comparison with a laboratory strain K-12
DNA Res. 8 (1), 11-22 (2001)
21156231
5 (bases 1 to 257071)
Ohnishi,M., Kurokawa,K., Makino,K., Yasunaga,T., Shinagawa,H. and
Hayashi,T.
Direct Submission
Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome
Information Research Center, 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail:ken-gen-info.osaka-u.ac.jp,
URL:http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365,
Fax:81-6-6879-2047)
genome project.

COMMENT
FEATURES
source
location/Qualifiers
1..257071
/organism="Escherichia coli O157:H7"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:83334"
93..1409
/gene="ECS2298"
93..1409
/gene="ECS2298"
/note="similar to B1592_ECOLI g1178786 percent identity
99 in 438 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative chloride channel"
/protein_id="BAB35721.1"
/db_xref="GI:13361765"
/translation="MFHCPFLMHHLHTYPRDLRTMFRLLIATVVGILAAFAVAFGRH
AMLEMLFLNDGSLVNAATRLSPMRLLIPALGGLAGLILMGMOKEPTOORPAP
PDYMEALQDGFQDVAASYKSLASLVYSSAIGREGAMITLALASCRQRPRT
ROBKMLWICGAGAAAYKAPLASLTAEVLEFTMLASGPVITSAVAALVSN
LHNSDALISVOLSTVQARVYALIISVGLAGLGPILLTLMNACHGFVSLKAP
PMQLAGGLIVGLISLFTPAVWNGSYVQSFLTAPPLMIITAGIPLCAVLAASSG
SGAPGVPTPLTIGALIGMLYGRSLGIMFPDGEITLLGLGMATLLAATTHADIM
SLTMCETMGEVOLPGLIACVIAVISRTLRHSIYRHTAQS"

gene
CDS
complement(1362..2069)
/gene="ECS2298"
complement(1362..2069)
/note="similar to B1593_ECOLI g1178787 percent identity
99 in 235 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB35722.1"
/db_xref="GI:13361766"
/translation="MELVLMKRFITGPTSVGKTVVSRALLQALASQKTVAGYRPV
AKGSETPGLRNKDALVLOSSTLELPYEAANVLAISEESSVASCINTLLISNG
LANLREKVHVVEGTGWRSLNDLSEWVQDLPVLAVYGIOECINHALITA
QALANDGLIGVANRINPGLAHYAEIIDVIGKKLPAPLIGLEPYLPRAEQRELGOY
IRLAMDRLVLAIDRVTV"

gene
CDS
complement(2182..3402)
/gene="ECS2300"
complement(2182..3402)
/note="similar to MLC_ECOLI g1178788 percent identity 99
in 406 aa (Conserved in E.coli K-12)"
/codon_start=1

/evidence=not_experimental
/transl_table=11
/product="putative NACC-like transcriptional regulator"
/protein_id="BAB35723.1"
/db_xref="GI:13361767"
/translation="MVAENQPHIDQIKOTNAGAVYRLIDQPSVRIDLSRLAOLAP
ASTRTKVRMELRAHLYDELEIKAGRGKRPNGVYETAMNHTLSIRISRGFELAR
DISSKIVVESEELAKDPSLDRLITSHIDQFIRHQRLEKRLTSIATLPEITDE
NGIVHMRPYEDVKEKMPLEALEQHGVPVYIQHDISAMTMAELFASRGADVQV
VIDHWAGVITFDGHLHAGSSSLVEIGHTOVDPYKRCYCGNGCLETIASVSTLE
LAOLRNAGVSMHGOPLTVSDICQALRGDLADITGVCAHYGRITAIWVNFN
PQKILISPLSKAADILFPVIDSIRQALPAYSCHISVESYQFSNQMAGALVKD
MAYNGSLIRLIQG"
complement(3537..4430)
/gene="ECS2301"
complement(3537..4430)
/note="similar to YNPL_ECOLI g1178789 percent identity
99 in 297 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative transcriptional regulator LYSR-type"
/protein_id="BAB35724.1"
/db_xref="GI:13361768"
/translation="MNIELRLRYFAVAEBLHFGAARLINSIPDSQIOALEEQ
IGARLARTNRSVLYLTAAGKOPADSRQILSWDDAARAERLHQEAGELRGFTSS
APFIRAVSDTSLSTRDYDPVHLQTRKEMTREQIADLIGTIDMGLRMTALPEYTH
AVIVHEPLAMTIPHDHPLANNPVTLAELAKEFEVFPDVGIVGIVDITGLMRVHL
TPVINGEAMTITIGLSAGISLTPASFRVQVLEKRWPIAEDADVSEKMLVWP
KHHEQSPARNRRIHLNLR"
4537..5790
/gene="ECS2302"
4537..5790
/note="similar to YNFM_ECOLI g1178780 percent identity
99 in 417 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative transport protein"
/protein_id="BAB35725.1"
/db_xref="GI:13361769"
/translation="MSRTTVVDAQPSDIDKQISQPNQFIKRTYQPMRYTLALFSA
GLATEFALICVOPILPVLSQEGFLTPANSISISITAMALIGLITGRLSDAIGKP
VMVATALLASICTLSTMTSMGIIIMRALIGLSGVAAGMIVLSEIHPSFVAF
SMGLYSIGNSIGMSGRLSGVYTFDFPNRIALAAIGCFALSAALMFMKILPESRHR
PSLIRKSLFINRLHMRGRPLFAEGPLMGSPVTLFNTIGYRLMSPHWQAV
VGLISLAYLTGTMSSPKACTMTGRGRCGYMLFSCVMLFGILMTLFSSIMLIFAGML
LPSAGFFAAHSAVSSWIGPRAKRAGQASSLYLFTYLGSSIAGTIGVFWMTGNG
VGAFLMLLVIALVGLRHLRLHA"
6187..6522
/gene="ECS2303"
6187..6522
/note="similar to ASR_ECOLI g1178781 percent identity 99
in 111 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="acid shock protein"
/protein_id="BAB35726.1"
/db_xref="GI:13361770"
/translation="MKKQIEGTMKKVLLVUVAAGLSAFAETATTPATATT
KAAPRATNHNKKQNKAAPOAKAKHNKNTKAEQKAPDEQAKQAKKNAKKSQQP
AKPAQAPAA"
6798..7322
/gene="ECS2304"
6798..7322
/note="similar to B1598_ECOLI g1178782 percent identity
94 in 171 aa but differs at C-ter (Conserved in E.coli
K-12)"

```
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB35727.1"
/db_xref="GI:1361771"
/translation="MRTIIVAVIGLISLTSFVFPADKPDVAKSANDEVSTLFFGHDR
VPSNKLWRYDIDHIDEGRDVPLGKRLKADGDGDIYPPAAPPDFELIVLRNPSSGITT
LPIREIKPRLEPH"
complement(7657..7986)
/gene="ECS2305"
complement(7657..7986)
/gene="ECS2305"
complement(7657..7986)
/translation="similar to B1599_ECOLI g1178783 percent identity
100 in 109 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence-not_experimental

Query Match      93.5%; Score 449.8; DB 1; Length 257071;
Best Local Similarity 97.3%; Pred. No. 2.3e-107;
Matches 468; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1 TGGCTCTCATTTGAATGATGATGCTTATTTGACACTATCTTTACCCAGCTCAACAGTTT 60
    |||||||
Db 38981 TGGCTCTCATTTGAATGATGATGCTTATTTGACACTATCTTTACCCAGCTCAACAGTTT 39040

QY 61 AATACCTGCGCAACATAAGGATGTTGTTAACTTAAGTCAAAAAAATAGCAATTTTC 120
    |||||||
Db 39041 AATACCTGCGCAACATAAGGATGTTGTTAACTTAAGTCAAAAAAATAGCAATTTTC 39100

QY 121 CAAGCAACAAAGCTAAATATGCAAAACCCAGTAAATAATCTGCTGAGCTATTATT 180
    |||||||
Db 39101 CAAGCAACAAAGCTAAATATGCAAAACCCAGTAAATAATCTGCTGAGCTATTATT 39159

QY 181 GCTAAGTACATTTACCCCTGGAAGTTAATGATCAATCAAGAGATGCGCTGTAAT 240
    |||||||
Db 39160 GCTAAGTACATTTACCCCTGGAAGTTAATGATCAATCAAGAGATGCGCTGTAAT 39219

QY 241 GAATCGTCTTATTTGAATTAACAGTTGATGCTTCTGCTGCTTCACTTCTTTG 300
    |||||||
Db 39220 GAATCGTCTTATTTGAATTAACAGTTGATGCTTCTGCTGCTTCACTTCTTTG 39279

QY 301 CGTGGCAGTACATTTGACAACTATGACCACTGAAACAGATGCTTCCGTACAAACAA 360
    |||||||
Db 39280 CGTGGCAGTACATTTGACAACTATGACCACTGAAACAGATGCTTCCGTACAAACAA 39339

QY 361 GTAAGCTCTGCACTTGTGAGCAGATGCTGCCGCTCGGGGTGCAGTGTTCACCTGTCG 420
    |||||||
Db 39340 GTAAGCTCTGCACTTGTGAGCAGATGCTGCCGCTCGGGGTGCAGTGTTCACCTGTCG 39399

QY 421 GATATTAAACAGGAATTTATCTTGTGTCGATGTTGGTGGATGTCAGAGGATGTA 480
    |||||||
Db 39400 GATATTAAACAGGAATTTATCTTGTGTCGATGTTGGTGGATGTCAGAGGATGTA 39459

QY 481 A 481
Db 39460 A 39460

RESULT 8
AF219227 358 bp DNA linear BCT 09-AUG-2000
LOCUS Escherichia coli beta-lactam resistance protein (blr) gene,
DEFINITION complete cds.
ACCESSION AF219227
VERSION AF219227.1 GI:8953940
KEYWORDS Escherichia coli.
SOURCE Escherichia coli.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia
REFERENCE 1 (bases 1 to 358)
```

```
AUTHORS Wong, R.S., McMurtry, L.M. and Levy, S.B.
TITLE 'Enterogenic' blr gene in Escherichia coli encodes a 41-residue
        membrane protein affecting intrinsic susceptibility to certain
        inhibitors of peptidoglycan synthesis
JOURNAL Mol. Microbiol. 37 (2), 364-370 (2000)
MEDLINE 20392465
PUBMED 10931331
REFERENCE 2 (bases 1 to 358)
AUTHORS Wong, R.S.Y., McMurtry, L.M. and Levy, S.B.
TITLE Direct Submission
        Submitted (27-DEC-1999) Molecular Biology & Microbiology, Tufts
        University School of Medicine, 136 Harrison Ave., Boston, MA 02111,
        USA
FEATURES
source 1..358
        location/Qualifiers
        /organism="Escherichia coli"
        /sub_species="K12"
        /db_xref="taxon:562"
        /map="min 36.6"
        93..>257
        /gene="blr"
        /note="similar to corresponding region of Escherichia coli
        sequences deposited in AB00258 and D90870"
        93..>257
        /gene="blr"
        /product="beta-lactam resistance protein"
        132..257
        /gene="blr"
        /note="new annotation of protein Blr"
        /codon_start=1
        /transl_table=11
        /product="beta-lactam resistance protein"
        /protein_id="AAF82191.1"
        /db_xref="GI:8953941"
        /translation="MNRILELTMGLVIVVSYLLGVASHIDNQPEQSASVQHK"

BASE COUNT 100 a 71 c 78 g 109 t
ORIGIN

Query Match      72.1%; Score 347; DB 1; Length 358;
Best Local Similarity 99.7%; Pred. No. 2.6e-80;
Matches 358; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 107 AATAGCAATTTTCCAAACGAAAGCTTAATATGCGCAAAAACCTCAATAAAATCTTG 166
    |||||||
Db 1 AATAGCAATTTTCCAAACGAAAGCTTAATATGCGCAAAAACCTCAATAAAATCTTG 59

QY 167 CTGAGCTATTATTTGTAAGTATACATTTACCCCTGAAGTTAATGATCAATCAAGAGAG 226
    |||||||
Db 60 CTGAGCTATTATTTGTAAGTATACATTTACCCCTGAAGTTAATGATCAATCAAGAGAG 119

QY 227 ATGTGGCTGTATGAATCGTCTTATTTGAATTAACAGTTGGATCGTCTTGTGTTTCA 286
    |||||||
Db 120 ATGTGGCTGTATGAATCGTCTTATTTGAATTAACAGTTGGATCGTCTTGTGTTTCA 179

QY 287 GTCAATCTTGTGGCGTGGGAGTCAATGAACTATCAGCCACTGAAAGAGATGCT 346
    |||||||
Db 180 GTCAATCTTGTGGCGTGGGAGTCAATGAACTATCAGCCACTGAAAGAGATGCT 239

QY 347 TCGGTACAAACAGTAAGCTCTGCACTTGTGAGAGCAGATGCTGCCGTCCGGGTGCAT 406
    |||||||
Db 240 TCGGTACAAACAGTAAGCTCTGCACTTGTGAGAGCAGATGCTGCCGTCCGGGTGCAT 299

QY 407 GTTTTCACTTGTGGATATTAAACGAGAAATTTATATCTTGTTCGATGTTGTGGTGA 465
    |||||||
Db 300 GTTTTCACTTGTGGATATTAAACGAGAAATTTATATCTTGTTCGATGTTGTGGTGA 358

RESULT 9
D90805 19211 bp DNA linear BCT 29-MAY-1997
LOCUS E.coli genomic DNA, Kohara clone #314(36.3-36.7 min.).
DEFINITION D90805 AB001340
ACCESSION D90805.1 GI:1742665
```

KEYWORDS Complete and shotgun sequencing; add: funa; func; hdb; hsdh; mali; maly; man; pmi; uida; gusa; gura; uibb; gusb; uidp; yihf.

SOURCE Escherichia coli (strain:K12) DNA, clone_lib:Kohara lambda miniset library clone:Kohara clone #314.

ORGANISM Escherichia coli

REFERENCE Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (sites)

AUTHORS Aiba,H., Baba,T., Fujita,K., Hayashi,K., Inada,T., Isono,K., Itoh,T., Kasai,H., Kashimoto,K., Kimura,S., Kitagawa,M., Kitagawa,M., Makino,K., Miki,T., Mizobuchi,K., Mori,H., Mori,T., Motomura,K., Nakade,S., Nakamura,Y., Nishio,Y., Oshima,T., Tagami,H., Takeda,J., Takemoto,K., Takeuchi,Y., Wada,C., Yamamoto,Y. and Horiuchi,T.

TITLE A 570-Kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 28.0-40.1 min region on the linkage map

JOURNAL DNA Res. 3 (6), 363-377 (1996)

MEDLINE 97251357

REFERENCE 2 (sites)

AUTHORS Aiba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitagawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T., Saito,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y. and Yano,M.

TITLE The systematic sequencing of the Escherichia coli genome in Japan

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 19211)

AUTHORS Mori,H.

TITLE Direct Submission

JOURNAL Submitted (14-DEC-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5

REFERENCE Takayama, Ikoma, Nara 630-01, Japan

TITLE (E-mail:hmori@gtc.aist-nara.ac.jp, Tel:81-7437-2-5660, Fax:81-7437-2-5669)

JOURNAL Project: Collaboration Information:

COMMENT The Japan E.coli genome DNA sequencing project

GROUP: The Japan E.coli genome DNA sequencing group

Members: (1995.4 - 1996.3)

Aiba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitagawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T., Saito,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y. and Yano,M.

Headed by: Name: Takashi Horiuchi

Address: National Institute of Basic Biology, Okazaki, 444, Japan

E-mail: kishori@nibb.ac.jp

Information operator: Name: Hirotsada Mori

Address: NARA Institute of Science and Technology, Ikoma, 630-01, Japan

E-mail: hmori@gtc.aist-nara.ac.jp

URL: The Japan E. coli genome database

http: bsw3.aist-nara.ac.jp.

Location/Qualifiers

1. 19211

/organism="Escherichia coli"

/strain="K12"

/db_xref="taxon:562"

/map="36.3 min"

/clone="Kohara clone #314"

/clone_lib="Kohara lambda miniset library"

/note="Nucleotide position 1688148-1707358 from the initiation site of ThnA (0 min.).-This clone is from

gene Kohara lambda miniset library"

complement(1..1272)

/gene="fuma"

complement(<1..1272)

/gene="fuma"

/note="ORF ID:0312#11

similar to [PIR Accession Number S07138]"

/codon_start=1

/transl_table=11

/product="Fumarate hydratase (EC 4.2.1.2) FumC"

/protein_id="BAAL5363.1"

/db_xref="GI:1742666"

/translation="MNTVSEKDSKALIDVPADKLMGATQTSLEFRISTEKPTSL IHALATKRAAKVNEEDUGLSEERKASATROADEYLACQHDDEFLATWQSGTOS NNMNEVLNARSELGVRGEMERKVPNDVNSOSSNDVPEPTAMHVAALLAKROL IQPLTITQTLNEKSRADFIVKIGRTHTDADATPLTGOEISGVAMLEHNLKIEYS LPHVAELAGTAVGTLNTHPEYARVADLAIVTICAPFVAPNFEALATCDALVQ AHGALKGLAASIMAKTANDVRLWASGRCGIGETISIPENRGSSIMPGKVNPCQCALT MCOCVMGNDVAIIMGASGNEFLNFRPMVTHNFIQSVRLADQNESRNKCAVGE PNRERINOLNLSMLVTLNTHIGDKAAET"

complement(1415..3061)

/gene="fuma"

complement(1415..3061)

/gene="fuma"

/note="ORF ID:0312#12

similar to [PIR Accession Number A03531]"

/codon_start=1

/transl_table=11

/product="Fumarate hydratase (EC 4.2.1.2) Fuma,

iron-dependent"

/protein_id="BAAL5364.1"

/db_xref="GI:1742667"

/translation="MSNRPFHVAQPEPLKKDTEYLLTSEHVSESEFEGQELIKVAP EALLTLARFHDASFMLRPAHQOVDILRPEASENRYVALQPLRSDIANAKVLP PCQDGTGTAIVGKRGQVMTGGDEAALARGYNTIENDLRSNAPLDMKEVNT GTNLPADQILVVDGDEYKELCIAKGGSANKTYLVOERKALLTPGLKLNVEKERT LGTAPAPVPIHIAFVIGGTSAETNLKVKLASAKTYDELPEQNGECAFARDELPEKEL LLEAONLIGAGQPGGKTFPHADIRVITLPHRGASCPYMGVSGSADNRITAKIRGQIW IERLEHNPGRVTPERLRKAGEAVAVDNLNRPKETLALQDSOTPVSTRLSNGTIIIG RDIHAHKLKERMDNGELGPQYIKDHPITYAGPAKTEGYASGSLGPTGTGRDSYVD LGAQGSMTIMLAKNGRSQOVTPACKKHGFYLGSIQPAVAVIAGSISIKSLECEYPEL GMEAIWKIEVEPFPATILVDKNGDNDFQQLQTLQCRVK"

3260..4435

/gene="mana, pmi"

3260..4435

/note="ORF ID:0312#13

similar to [Swissprot Accession Number P00946]"

/codon_start=1

/transl_table=11

/product="Mannose-6-phosphate isomerase (EC 5.3.1.8)

(Phosphomannose isomerase) (PMI) (Phosphohexomutase)."

/protein_id="BAAL5365.1"

/db_xref="GI:1742668"

/translation="MOKLINSQNVYAMGSKTALTEYGMENPSSQPMALMCAHNRKS SSRVQNAAGDIVSLRDVTSDDSTLGEVAVKRGELPLFLKYLCAQPSLSTOVHNRK HNSEIGFAENAGAPMDAERNYKDPNKRPELVATLPEFLANARRESEVSLIQR VAGHAPAIHFLQDPDAERLSEFASILMMQGEKSRALAILKSLADSOOGFEWQIR LLSFEPDSEGLFSPLLNVVKNLPGEAHLFEETGHAYLQGVALLVMSNDVRLAG LIPKYIDIPELIVANVKEFAKPPANQILLTOPVKQAEIDPPIVDDFAFSLHDSKRETT ISQGSAAITFCVBDGATLWKGSQQLDKGESAFLIANESPVTYVKGHRLARYNKL"

4536..6044

/gene="yihf"

4536..6044

/note="ORF ID:0312#14

similar to [Swissprot Accession Number P32128]"

/codon_start=1

/transl_table=11

/protein_id="BAAL5366.1"

/db_xref="GI:1742669"

/translation="MNKSLVAVGVIVYALGVVYMGAWYTGKRIETHLEDMVQAANQL KITAPESNLEVSQNTNRGVFSQQLDLVKKPIAGKNEPWIKSGSVIFNESVDHGFP

LAOLKRLIPSMASIQTLVNNVSKPLFPMANGEPFELINSRIGSGSSDSISLK
PLNVEOKDEKVAFGSGEPOLNADRDGKAISLSGASGRIDAVNEYKQVLTENNLK
TDGSETLASFGERVGNKLTLEKMTISVEKELALLEMEISGSDIADPEKNTINSOL
DYSLSLKVONODGSKLTUKGOIDEAMHOSQOYNAQTOALLAPPEANPELVE
QEKYAEFSAFALPMLKGDPAVITTAIRKSGMSOGSGLSLFLKDPPTTEAQTOLA
QEVDRSVASLDKLTIPYDAMATERTQVAKIEGQEOADAKLAKOVSASAMQMFR
LTTTODNTTTSLOIANGQITLNGOKSLDEFGVGFAPALNVAVPAIPQO"

gene
complement(7574. .8947)
/gene="uidb. gusb. uidp"
complement(7574. .8947)
/gene="uidb. gusb. uidp"
/note="ORF_ID:0314#2
similar to [SwissProt Accession Number P30868]"
/codon_start=1
/transl_table=11
/product="Glucuronide carrier protein (Glucuronide
permease)."

/db_xref="GI:1742670"
/translation="NQQLSWRTIVGSLGDVNAFAMGALFILSYTVDVAGVAA
AAGTMLLVRFADADVPAGRVDSVNTRWCKRPFLLFGTAPLMITSVLVFWLTD
WSHGRVYVAYLTVMGLCYSLNIPYGLATATQOPQSHARGAARGIAASLTFV
CLAFIGSPKSNKSPSEWVSVEHWTIVYALAGMVLFECEKSTRENNVRIYAQSLN
ISLOTLRNRPLFMCIGALCYLSTFPAVSASLFEYVRVNDTGLFTVLVYQMLG
TVASAPVPGMVARIGKNTPLIGALIGCGLFEFWSVMSLPALVALAISIGOG
VMYVMALADPTVEXGELTGVRIGELTYSLFSTRKCGAGIGSIDAFILGSGYI
ANQVOTPEYIKIRTSIALVPCGFMFLAFVIITFPLDKRKEIVEIDNNKKVQOO
LSDITN"

gene
complement(8944. .10755)
/gene="uida. gusa. gura"
complement(8944. .10755)
/gene="uida. gusa. gura"
/note="ORF_ID:0314#3
similar to [SwissProt Accession Number P05804]"
/codon_start=1
/transl_table=11
/product="b-glucuronidase (EC 3.2.1.31) (GUS)
(b-d-glucuronoside glucuronosylhydrolase)."

/protein_id="BA15368.1"
/db_xref="GI:1742671"
/translation="MLRPVETPREIKRIDGIMAFSLDRENCIDIORWESALQESRA
IAPGSFNDQFADADIRNAGVNWTOREFFIRGMAGQIVLRFDVYTHYGRVWANO
EYMEHGGATPEADVTPIYIAGKSVIRIVCVNNELMWTIPGMVITDENKRSQSY
FHDFFNYAGIHSVMLYTPNTWVDITVTVADCNASVDQWQVANGDVSELVD
ADQOVVATGCGTSGTLOVNPMLDQEGEYIELCYTAKSOTECIDILPRLVGRISYAV
KGRQFLINHKPRYFTGSEGRHEDADLRGKGDVNLVHDALMDIGANSYRSHRYA
EKMLADKHPHSYVMSIANEPPTROGARRETFAPLAKYTRKIDPRPTTCYNNMCD
KELIARDKHPHSYVMSIANEPPTROGARRETFAPLAKYTRKIDPRPTTCYNNMCD
AHTDTISLDELVCILNRYIGWYVQSDLETAELVLEKELIAMQERKLHPTITTEGYD
TLAHLHSMYTDMMSEYQCAMLDMYHRVDFRVSAYVGEQVWMEFADPTISGILRVGNN

Query Match 42.2%; Score 203; Db 1; Length 19211;
Best Local Similarity 99.5%; Pred. No. 1.1e-42;
Matches 214; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 TGCCTCTCATTTGAAGTATGATGCGCTATTGACACTATCTTTACCCACGCTCAACAGTTT 60
|||||
Db 18998 TGCCTCATATTGAAGTATGATGCGCTATTGACACTATCTTTACCCACGCTCAACAGTTT 19057
QY 61 AATAAAGTCCAGCAATTAAGGATGCTGTTAACTTAAGCAAAAAAATAGCGAATTTTC 120
|||||
Db 19058 AATAAAGTCCAGCAATTAAGGATGCTGTTAACTTAAGCAAAAAAATAGCGAATTTTC 19117
QY 121 CAACGAACAAAAGCTAATATATCGCAAAACCTAGTAAAAATCTTGCCTGAGCATATTATT 180
|||||
Db 19118 CAACG-ACAAAAGCTAATATATCGCAAAACCTAGTAAAAATCTTGCCTGAGCATATTATT 19176
QY 181 GCTAAGTAACTTTACCCCTGAAGTTAATGATC 215
|||||
Db 19177 GCTAAGTAACTTTACCCCTGAAGTTAATGATC 19211

RESULT 10

AE008763/c
LOCUS AE008763 23009 bp DNA linear BCF 31-JUL-2002
DEFINITION Salmoneilla typhimurium LT2, section 67 of 220 of the complete
genome.
ACCESSION AE008763 AE006468
VERSION AE008763.1 GI:16419969
KEYWORDS
SOURCE Salmoneilla typhimurium LT2.
ORGANISM Salmoneilla typhimurium LT2.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmoneilla.
REFERENCE 1 (bases 1 to 23009)
McCllland,M., Sanderson,R.E., Spieth,D., Clifton,S.W.,
Latreille,P., Courtney,L., Porwollik,S., Ali,J., Dante,M., Du,F.,
Hou,S., Layman,D., Leonard,S., Nguyen,C., Scott,K., Holmes,A.,
Grewal,N., Mulvaney,E., Ryan,E., Sun,H., Flores,L., Miller,W.,
Stonking,T., Nhan,M., Waterston,R. and Wilson,R.K.
Complete genome sequence of Salmoneilla enterica serovar Typhimurium
LT2
JOURNAL Nature 413 (6858), 852-856 (2001)
MEDLINE 21534948
PUBMED 11677609
REFERENCE 2 (bases 1 to 23009)
The Salmoneilla typhimurium Genome Sequencing Project.
Direct Submission
Submitted (29-MAR-2001) Genome Sequencing Center, Department of
Genetics, Washington University School of Medicine, 4444 Forest
Park Boulevard, St. Louis, MO 63108, USA
COMMENT Supported by NIH grant 5U 01 AI43283

COMMENT

Coding sequences below are predicted from manually evaluated
computer analysis, using similarity information and the programs;
GLIMMER: <http://www.tigr.org/softlab/glimmer/glimmer.html> and
GeneMark: <http://opal.biology.gatech.edu/geneMark/>

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto
Encyclopedia of Genes and Genomes: <http://www.genome.ad.jp/kegg/>,
and Pedro Romero and Peter Karp at EcoCyc:
<http://ecocyc.PangeaSystems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites
were kindly provided by Heladia Salgado, Julio Collado-Valdes and
ReguionDB:
http://kinich.cifn.unam.mx:8650/db/reguiondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one m3 subclone.

FEATURES

source

1..23009
/organism="Salmoneilla typhimurium LT2"
/strain="LT2; SCSG 700720"
/db_xref="ATCC:700720"
/db_xref="taxon:99287"
/note="LT2"

gene

88..1374
/gene="tyrS"
/note="STM1449"

RBS

88..93
/gene="tyrS"
/note="putative RBS for tyrS. ReguionDB:STMS1H001673"
100..1374
/gene="tyrS"
/BC_number="6.1.1.1"

CDS

/note="similar to E. coli tyrosine tRNA synthetase
(AAC74709.1); Blastp hit to AAC74709.1 (424 aa), 96%
identity in aa 1 - 424"
/codon_start=1
/transl_table=11
/product="tyrosine tRNA synthetase"

RESULT 11	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AL627271	233050 bp	DNA	linear	BC1 06-JUN-2002			
AL627271		Salmonella enterica serovar Typh1 (Salmonella typh1) strain CT18,					
AL627271	AL513382	Complete chromosome, segment 7/20.					
AL627271.1	GI:16502733						
REFERENCE							
AUTHORS							
1 (bases 1 to 233050)							
Parkhill,J., Dougan,G., James,K.D., Thomson,N.R., Pickard,D.,							
Wain,J., Churcher,C., Mungall,K.L., Bentley,S.D., Holden,M.T.G.,							
Sobralha,M., Baker,S., Basham,D., Brooks,K., Chillingworth,T.,							
Conerton,P., Cronin,A., Davis,P., Davies,R.M., Dowd,L., White,N.,							
Farrar,J., Feltham,T., Hamlin,N., Heaugh,A., Hien,T.T., Holroyd,S.,							
Jagels,K., Krogh,A., Larsen,T.S., Leather,S., Moulé,S., O'Gaora,P.,							
Perry,C., Quail,M., Rutherford,K., Simmonds,M., Skelton,J.,							
Stevens,K., Whitehead,S. and Barrrell,B.G.							
entire genome sequence of a multiple drug resistant Salmonella							
enterica serovar Typh1 CT18							
Nature 413 (6858), 848-852 (2001)							
2 (bases 1 to 233050)							
Parkhill,J.							
Direct Submission							
Submitted (25-OCT-2001) Submitted on behalf of the Salmonella							
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,							
Hinxton, Cambridge CB10 1SA, UK							
E-mail: parkhill@sanger.ac.uk							
Notes:							
Details of S. typh1 sequencing at the Sanger Centre are available							
on the World Wide Web.							
(URL, http://www.sanger.ac.uk/Projects/S_typh1/).							
FEATURES							
source							
1..233050							
/organism="Salmonella enterica subsp. enterica serovar							
Typh1"							
/strain="CT18"							
/db_xref="taxon:90370"							
199..1581							
/gene="STY1645"							
199..1581							
/gene="STY1645"							
/note="Similar to Pseudomonas aeruginosa							
arginine/ornithine antiporter arcd SW:ARCD_PSEAE (P18275)							
(482 aa) fasta scores: E(): 0, 45..5% id in 479 aa							
fasta hit to YJDE_ECOLI (445 aa), 31% identity in 450 aa							
overlap							
Orthologue of E. coli ARCD_ECOLI: Fasta hit to ARCD_ECOLI							
(460 aa), 93% identity in 460 aa overlap"							
/codon_start=1							
/transl_table=1							
/product="putative amino acid permease"							
/protein_id="CAD01890.1"							
/db_xref="GI:16502733"							
/db_xref="SPTREMBL:O8XG1"							
/translation="MEKIGLSALTALVLSMIGAGVPSLPONMAAVASPAALLIGNAN							
ITGGILAFAMLLIRIRSELDGGITTYTAREGFGELIGSCSMGWLCVAVINVS							
LIVAFSAFTPTDPALRLFDGNTWQALIVASVLLVHVLIRGVOTASIVLVAT							
LAKLIPGAFIVLIMFKEDTFTLDFEGLIGIPWEQVNTMLITLWEIIGEGAN							
VYSAAKNRDVGRAITLAVLAGIYLLVLTLLSGVAPBELAKMRPSSMGALMKV							
MGSPKEIIAAGLIYSCGAVLAWMTIAAEVPLAAAKSPGFARONAGASAPS							
WLNRCVQICVLTIWTPSDYNTLTITASEMTIIVPEYLGAFILKIATRPILHRAVGIG							
ACTGATMLVYASGPMHLLVSVLTAIPGLVFLVIRTKHKDNTVLRQDEVYIIGLLVPA							
AVPAPMLVWG*							

```

misc_feature
199. .1563
/gene="STY1645"
/note="Pfam match to entry PF00324_aa_permeases, AmIno
acid permease, score .290.50, E-value 0.062"
complement(1643. .1978)
/gene="STY1646"
complement(1643. .1978)
/gene="STY1646"
/note="orthologue of E. coli YDCG_ECOLI: Fasta hit to
YDCG_ECOLI (111 aa), 81% identity in 111 aa overlap"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAD01891.1"
/db_xref="GI:16502735"
/db_xref="SPTRMBL:082652"
/translation="MGIVIKALGALNVVLIGLSTKKNYIAGLIPLEPFTALAHY
IVASEGIDAMNTTIVFSMRSIIPFIYLAIVLFSGVMPLPALGCAVYCCMLSNML
LIFCWLKMH"
2105. .2836
/gene="STY1647"
/note="rsta"
2105. .2836
/gene="STY1647"
/note="similar to Staphylococcus aureus response regulator
YJCF yycF TR:Q9XCM7 (EMBL:AF136709) (233 aa) fasta scores:
E1): 8.8e-28, 37.6% id in 224 aa
Fasta hit to TORR_ECOLI (230 aa), 31% identity in 213 aa
overlap
Fasta hit to BAER_ECOLI (240 aa), 34% identity in 229 aa
overlap
Fasta hit to YLCA_ECOLI (227 aa), 32% identity in 232 aa
overlap
Fasta hit to YEDW_ECOLI (239 aa), 34% identity in 233 aa
overlap
Fasta hit to OMPR_ECOLI (239 aa), 33% identity in 235 aa
overlap
Fasta hit to CPXR_ECOLI (232 aa), 37% identity in 233 aa
overlap
orthologue of E. coli rsta (RSTA_ECOLI); Fasta hit to
RSTA_ECOLI (242 aa), 85% identity in 239 aa overlap"
/codon_start=1
/transl_table=11
/product="putative two-component response regulator"
/protein_id="CAD01892.1"
/db_xref="GI:16502736"
/db_xref="SPTRMBL:082651"
/translation="MNRIVFEDDAEVGSLIAAYIAKHDIYIEPRDRAEDILIT
QPVVLVIDIMLPKDGDMITCRDLRRHMGPIYLLTSDSDNMHILALEMGCDVILIT
TPPAVLAIRLRILRQSEQTOAKSLQESALPHKALFGALTIDPLRAVAGLGFIT
STLTAFFELLMELATAGQIMDRDALTLIRGVNYDGLDRSDVAISIRLRKLLDSAA
EPRIRITKNGYGLFAPHAMDETGT"
2108. .2440
/gene="STY1647"
/note="Pfam match to entry PF00072_response_reg_Response
regulator receiver domain, score 62.70, E-value 7.9e-15"
2582. .2794
/gene="STY1647"
/note="Pfam match to entry PF00486_trans_reg_C,
transcriptional regulatory protein, C terminal, score
62.90, E-value 1.8e-17"
complement(2901. .3260)
/gene="STY1648"
complement(2901. .3260)
/gene="STY1648"
/note="No significant database matches"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAD01893.1"
/db_xref="GI:16502737"
/db_xref="SPTRMBL:082650"
/translation="MNGRPHNLNIIISANANFSYMLLINTKVTIGNIAYIYGEIYCIYI

```


COMMENT

On Nov 10, 2000 this sequence version replaced gi:3802827.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center Project name: L10498

Center clone name: 137_P11

NOTE: This is a 'working draft' sequence. It currently consists of 52 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 164: contig of 164 bp in length
165 264: gap of 100 bp
265 848: contig of 584 bp in length
849 948: gap of 100 bp
949 1605: contig of 657 bp in length
1606 1705: gap of 100 bp
1706 2318: contig of 613 bp in length
2319 2418: gap of 100 bp
2419 3428: contig of 1010 bp in length
3429 3528: gap of 100 bp
3529 4191: contig of 663 bp in length
4192 4291: gap of 100 bp
4292 5090: contig of 799 bp in length
5091 5190: gap of 100 bp
5191 5963: contig of 773 bp in length
5964 6063: gap of 100 bp
6064 7369: contig of 1306 bp in length
7370 7469: gap of 100 bp
7470 8607: contig of 1138 bp in length
8608 8707: gap of 100 bp
8708 17424: contig of 8717 bp in length
17425 17524: gap of 100 bp
17525 18771: contig of 1247 bp in length
18772 18871: gap of 100 bp
18872 19924: contig of 1053 bp in length
19925 20024: gap of 100 bp
20025 21370: contig of 1346 bp in length
21371 21470: gap of 100 bp
21471 23256: contig of 1786 bp in length
23257 23356: gap of 100 bp
23357 25010: contig of 1654 bp in length
25011 25110: gap of 100 bp
25111 26876: contig of 1766 bp in length
26877 26976: gap of 100 bp
26977 28146: contig of 1170 bp in length
28147 28246: gap of 100 bp
28247 29546: contig of 1300 bp in length
29547 29646: gap of 100 bp
29647 31979: contig of 2333 bp in length
31980 32079: gap of 100 bp
32080 33806: contig of 1727 bp in length
33807 33906: gap of 100 bp
33907 35287: contig of 1381 bp in length
35288 35387: gap of 100 bp
35388 36754: contig of 1367 bp in length
36755 36854: gap of 100 bp
36855 38157: contig of 1303 bp in length
38158 38257: gap of 100 bp
38258 40144: contig of 1887 bp in length
40145 40244: gap of 100 bp
40245 42144: contig of 1900 bp in length
42145 42244: gap of 100 bp

42245 45014: contig of 2770 bp in length
45015 45114: gap of 100 bp
45115 47299: contig of 2185 bp in length
47300 47399: gap of 100 bp
47400 49714: contig of 2315 bp in length
49715 49814: gap of 100 bp
49815 51867: contig of 2053 bp in length
51868 51967: gap of 100 bp
51968 54711: contig of 2744 bp in length
54712 54811: gap of 100 bp
54812 57606: contig of 2795 bp in length
57607 57706: gap of 100 bp
57707 60300: contig of 2594 bp in length
60301 60400: gap of 100 bp
60401 62959: contig of 2559 bp in length
62960 63059: gap of 100 bp
63060 64544: contig of 1485 bp in length
64545 64644: gap of 100 bp
64645 66943: contig of 2299 bp in length
66944 67043: gap of 100 bp
67044 68675: contig of 1632 bp in length
68676 68775: gap of 100 bp
68776 71620: contig of 2845 bp in length
71621 71720: gap of 100 bp
71721 74270: contig of 2530 bp in length
74271 74370: gap of 100 bp
74371 78405: contig of 4035 bp in length
78406 78505: gap of 100 bp
78506 81246: contig of 2741 bp in length
81247 81346: gap of 100 bp
81347 83767: contig of 2421 bp in length
83768 83867: gap of 100 bp
83868 86612: contig of 2745 bp in length
86613 86712: gap of 100 bp
86713 89372: contig of 2660 bp in length
89373 89472: gap of 100 bp
89473 93415: contig of 3943 bp in length
93416 93513: gap of 100 bp
93514 98713: contig of 5198 bp in length
98714 98813: gap of 100 bp
98814 103055: contig of 4242 bp in length
103056 103155: gap of 100 bp
103156 109896: contig of 6841 bp in length
109897 110096: gap of 100 bp
110097 117716: contig of 7620 bp in length
117717 117816: gap of 100 bp
117817 124867: contig of 7051 bp in length
124868 124967: gap of 100 bp
124968 133690: contig of 8723 bp in length
133691 133790: gap of 100 bp
133791 134978: contig of 1188 bp in length.

FEATURES

source

1..134978

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP11-137P11"

/clone_lib="RPC1-11 Human Male BAC"

/note="assembly-fragment"

clone_end:SP6

vector_side:left"

265..848

/note="assembly-fragment"

misc-feature

Query Match

8.8%; Score 42.4; DB 2: Length 134978;

Best Local Similarity 53.7%; Pred. No. 1.2; Mismatches 88; Conservative 0; Indels 0; Gaps 0;

OY 26 ATTGACACTATCTTTTACCCAGCTCAACAGTTTAAATACCTGCGCAATAAGGATG 85
DB 3207 ATTGACTCTGCTCTCTCCGCGCTCATCTTCCTTATCCCTTCCTTCCTTACCTATG 3148

```

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (17-AUG-2000) Whitehead Institute/MIT Center for Genome
Research 320 Charles Street, Cambridge, MA 02141, USA
On Nov 10, 2000 this sequence version replaced g1:9857979.
All repeats were identified using RepeatMasker:
Smtt, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIMR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L10489
Center clone name: L11.G_3
-----
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 154868 bases at least Q40
Consensus quality: 159752 bases at least Q30
Consensus quality: 161420 bases at least Q20
Insert size: 155000; agarose-fp
Insert size: 162284; sum-of-contigs
Quality coverage: 4.4 in Q20 bases; agarose-fp
Quality coverage: 4.2 in Q20 bases; sum-of-contigs

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	*	1	2073:	contig of 2073 bp	in length
*	*	2074	2173:	gap of 100 bp	in length
*	*	2174	3577:	contig of 1404 bp	in length
*	*	3578	3677:	gap of 100 bp	in length
*	*	3678	4992:	contig of 1315 bp	in length
*	*	4993	5092:	gap of 100 bp	in length
*	*	5093	7181:	contig of 2089 bp	in length
*	*	7182	7281:	gap of 100 bp	in length
*	*	7282	9225:	contig of 1944 bp	in length
*	*	9226	9325:	gap of 100 bp	in length
*	*	9326	11777:	contig of 2452 bp	in length
*	*	11778	11877:	gap of 100 bp	in length
*	*	11878	14570:	contig of 2693 bp	in length
*	*	14571	14670:	gap of 100 bp	in length
*	*	17522	17522:	contig of 2852 bp	in length
*	*	17523	17622:	gap of 100 bp	in length
*	*	17623	20980:	contig of 3358 bp	in length
*	*	20981	21080:	gap of 100 bp	in length
*	*	21081	25542:	contig of 4462 bp	in length
*	*	25543	25642:	gap of 100 bp	in length
*	*	25643	31250:	contig of 5608 bp	in length
*	*	31251	31350:	gap of 100 bp	in length
*	*	31351	36798:	contig of 5448 bp	in length
*	*	36799	36898:	gap of 100 bp	in length
*	*	36899	43788:	contig of 6690 bp	in length
*	*	43789	43888:	gap of 100 bp	in length
*	*	43889	50473:	contig of 6585 bp	in length
*	*	50474	50573:	gap of 100 bp	in length
*	*	50574	59219:	contig of 8646 bp	in length
*	*	59220	59319:	gap of 100 bp	in length
*	*	59320	70858:	contig of 11539 bp	in length
*	*	70859	70958:	gap of 100 bp	in length
*	*	70959	107991:	contig of 37033 bp	in length
*	*	107992	108091:	gap of 100 bp	in length
*	*	108092	127851:	contig of 19760 bp	in length
*	*	127852	127951:	gap of 100 bp	in length
*	*	127952	153850:	contig of 25999 bp	in length
*	*	153951	154050:	gap of 100 bp	in length
*	*	154051	164184:	contig of 10134 bp	in length
Location/Qualifiers					
1..164184					
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
/chromosome="8"					
/map="8"					
/clone="R1b-11163"					
/clone.lib="RPC1-11 Human Male BAC"					
1..2073					
/note="assembly-fragment					
clone_end:SP6					
vector_side:left"					
2174..3577					
/note="assembly-fragment"					
3678..4992					
/note="assembly-fragment"					
5093..7181					
/note="assembly-fragment"					
7282..9225					
/note="assembly-fragment"					
9326..11777					
/note="assembly-fragment"					
11878..14570					
/note="assembly-fragment"					
14671..17522					

```

misc_feature      17623..20980      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      21081..25542      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      25643..31250      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      31351..36798      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      36899..43788      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      43889..50473      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      50574..59219      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      59320..70858      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      70959..107991      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      108092..127851      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      127952..153950      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      154051..164184      /note="assembly_fragment"
                    /note="assembly_fragment"
misc_feature      clone_end:T7
                    vector_side:right"
BASE COUNT      50016 a 29587 c 29331 g 53349 t 1901 others
ORIGIN

```

```

Query Match
Best Local Similarity 5.8%; Score 42.4; DB 2; Length 164184;
Matches 88; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

```

```

QY 26 ATTTGACACTATCTTTACCCAGCTCAACAGTTTAAATACCTGCCAGCATTAAGCATG 85
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 110406 ATTTGACTGCGCTCTTCCGCTCCATCTTATCCCTCTCTCTCTCTCTCTCTCTCT

```

```

QY 86 TTGTACTTAACTCAAAAAAATAGCGAATTTTCCACAGCAAAAGCTAATATGCGCA 145
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 110346 TGGCTATATCTTGAAACAACAAATACATACATAGCAACAGAAATATACATATMAAA 110287
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

QY 146 AAAACCTCAGTAAATCTTCTGCGAGCTATTTGCTAAGTAA 189
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 110286 ACAACTTCAGTAAATATCCAGCTAATATTTGATCTCCAGTGA 110243
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 14
AC099687      167577 bp      DNA      linear      PRI 12-JUN-2002
LOCUS      Homo sapiens chromosome 11, clone CTD-2507G9, complete sequence.
AC099687
AC099687.5      GI:21392510
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 167577)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE      Homo sapiens chromosome 11, clone CTD-2507G9
JOURNAL
PUBLISHED
2 (bases 1 to 167577)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campolano,A., Chang,J., Chazaro,B.,
Choepe,I., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,J., Lecoque,K.,
Lamazares,R., Lander,T., Lehoczy,J., Levine,R., Liu,G.,

```

```

TITLE
JOURNAL
Submitted (18-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 167577)
AUTHORS

```

```

Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,Y., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (08-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 167577)
AUTHORS

```

```

REFERENCE
JOURNAL
Submitted (12-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 12, 2002 this sequence version replaced gi:17386347.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RW/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

```

repeat_region	/organism="Homo sapiens"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="11"
repeat_region	/map="11"
repeat_region	/clone="C1D-2507G9"
repeat_region	/clone_id="C1TD2 Human BAC 40..165
repeat_region	/rpt_family="FLAM_C"
repeat_region	complement(216..670)
repeat_region	/rpt_family="LTR32"
repeat_region	883..1192
repeat_region	/rpt_family="Alusx"
repeat_region	2230..2449
repeat_region	/rpt_family="L1MC4a"
repeat_region	complement(2471..2688)
repeat_region	/rpt_family="MER51A"
repeat_region	complement(2717..2970)
repeat_region	/rpt_family="MER1A"
repeat_region	complement(3047..3095)
repeat_region	/rpt_family="MER51A"
repeat_region	3132..4333
repeat_region	/rpt_family="L1MC4a"
repeat_region	complement(4363..4956)
repeat_region	/rpt_family="MER50"
repeat_region	complement(4957..5262)
repeat_region	/rpt_family="AluY"
repeat_region	complement(5263..5330)
repeat_region	/rpt_family="MER50"
repeat_region	5337..6142
repeat_region	/rpt_family="L1MC4a"
repeat_region	6143..6441
repeat_region	/rpt_family="AluY"
repeat_region	6442..6513
repeat_region	/rpt_family="L1MC4a"
repeat_region	6519..6602
repeat_region	/rpt_family="(TA)n"
repeat_region	6604..6758
repeat_region	/rpt_family="LTR8A"
repeat_region	6759..6795
repeat_region	/rpt_family="MER1B"
repeat_region	6956..7078
repeat_region	/rpt_family="L1MC4"
repeat_region	complement(7702..8619)
repeat_region	/rpt_family="L1P1A16"
repeat_region	complement(9197..9374)
repeat_region	/rpt_family="L1M"
repeat_region	complement(10565..10788)
repeat_region	/rpt_family="MIR"
repeat_region	complement(10942..11305)
repeat_region	/rpt_family="L2"
repeat_region	11691..11880
repeat_region	/rpt_family="L2"
repeat_region	11892..11924
repeat_region	/rpt_family="(TGA)n"
repeat_region	12948..12972
repeat_region	/rpt_family="AT-rich"
repeat_region	13462..13752
repeat_region	/rpt_family="AluSx"
repeat_region	complement(13989..14140)
repeat_region	/rpt_family="MIR"
repeat_region	14464..15015
repeat_region	/rpt_family="MER7A"
repeat_region	15114..15682
repeat_region	/rpt_family="MIR"

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

Center project name: bm188a3
----- Summary Statistics -----
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator; 6% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 93% of reads
Consensus quality: 189901 bases at least Q40
Consensus quality: 189915 bases at least Q30
Consensus quality: 189915 bases at least Q20
Insert size: 189915; sum-of-contigs
Insert size: 188814; 4.0% error; agarose-fp
Quality coverage: 18.04x in Q20 bases; sum-of-contigs Quality
coverage: 20.10x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 11148: contig of 11148 bp in length
* 11149 11248: gap of 100 bp
* 111249 190015: contig of 78767 bp in length.
Location/Qualifiers
1.190015
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-188A3"
/clone_11b="RPCT-23"
1.11148
/note="assembly_fragment:07880.0"
111249.190015
/note="assembly_fragment:07880.1"
BASE COUNT 60196 a 40379 c 37523 g 51817 t 100 others
ORIGIN
misc_feature
1.11148
/note="assembly_fragment:07880.0"
misc_feature
111249.190015
/note="assembly_fragment:07880.1"
BASE COUNT 60196 a 40379 c 37523 g 51817 t 100 others
ORIGIN

Query Match 8.88; Score 42.2; DB 2; Length 190015;
Best Local Similarity 58.3%; Pred. No. 1.3;
Matches 74; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 103 AAAAATAGCAATTTTCCAAAGCAAAAGCTAAATATGCAAAAACCTCAGTAAAT 162
||||| ||| ||| ||||| ||||| ||| ||| |||
Db 21377 AAAAATAGCAATTTTCCAAAGCAAAAGCTAAATATGCAAAAACCTCAGTAAAT 21436

QY 163 CTGCTGAGCTATTATGCTAAGTAACATTTACCCCTGAAGTTAATGATCAATCAAG 222
||| | ||||| ||| ||| ||| ||| ||| ||| |||
Db 21437 ATTGTATTATGTTATGTTACATGACATTAACCATCTTTTCTAAGAGGAAATATAT 21496

QY 223 AGAGATG 229
|||||
Db 21497 AGAGCTG 21503

```

Search completed: February 7, 2003, 20:36:34
Job time : 2920 secs


```

FT mutation replace (235..237, TAG)
FT /*tag= e
FT /label= "Amber-mutation_V20
FT /note= "An amber mutation at this position does not
FT abolish beta-lactam resistance"
FT CDS 239..364
FT /*tag= f
FT /product= "BUR protein"
FT replace (247..249, TAG)
FT /*tag= g
FT /label= "Amber-mutation_L24
FT /note= "An amber mutation at this position abolishes
FT beta-lactam resistance, but this can be reversed
FT using an amber suppressor"
FT mutation replace (292..294, TAG)
FT /*tag= h
FT /label= "Amber-mutation_L39
FT /note= "An amber mutation at this position abolishes
FT beta-lactam resistance, but this can be reversed
FT using an amber suppressor"
FT primer_bind 332..355
FT /*tag= 1
FT /note= "Primer R3 (not given in the specification)"
FT 338..339
FT /*tag= 1
FT /note= "Insertion point for mini-TnphoA transposon in
FT E. coli strain RM583"
FT primer_bind 448..465
FT /*tag= k
FT /note= "Primer R1 (not given in the specification)"
FT 470..481
FT /*tag= 1
FT /label= "ORE71-promoter
FT /note= "Promoter for ORE71 (putative histone-like
FT negative regulator gene"
FT WO200177176-A2.
FT 18-OCT-2001.
FT 06-APR-2001; 2001WO-US11363.
FT 06-APR-2000; 2000US-195505P.
FT 14-JUL-2000; 2000US-218380P.
FT (TUFT ) TUFTS COLLEGE.
FT Levy SB, McMurry LM;
FT P-PSDB; AAM52607.
FT WPI: 2002-010898/01.
FT P-PSDB; AAM52607.
FT New membrane protein, designated Beta Lactam-356 polypeptides, that
FT affect susceptibility to antibiotics which affect peptidoglycan
FT synthesis in microbes, useful for identifying modulators for treating
FT infections _
FT
FT Claim 1, Fig 1A-B, 104pp: English.
XX
XX This sequence represents a beta-lactam resistance (blr) gene from
XX Escherichia coli. The blr gene encodes a 41 amino acid transmembrane
XX protein (also designated Beta Lactam-358 in the specification) with a
XX molecular weight of 4556 Daltons which confers resistance to beta-lactam
XX antibiotics. The blr gene was initially identified in a 602 bp
XX hypothetical intergenic region of the genome of an antibiotic-susceptible
XX strain of E. coli, RM583. In the RM583 strain, the blr gene is disrupted
XX by insertion of the transposon mini-TnphoA. A clone bearing an ORF (open
XX reading frame)-containing 358 bp segment of the "intergenic" region was
XX found to restore beta-lactam resistance in RM358. Certain amber
XX mutations introduced into the 358 bp clone were found to abolish
XX beta-lactam resistance and resistance could be restored using an
XX amber suppressor, indicating that a protein encoded by the clone was
XX responsible for mediating beta-lactam resistance. It is suggested that

```

```

CC the BUR protein may act by increasing a beta-lactamase activity in a
CC cell, possibly being a component of a uncharacterised membrane-bound
CC efflux pump capable of expelling beta-lactams from the periplasm. The
CC BUR protein may be used in methods to identify compounds that modulate
CC antibiotic resistance in a bacterium. The BUR protein, gene, BUR protein
CC homologues, BUR agonists or antagonists, and anti-BUR antibodies may be
CC used to treat bacterial infection, particularly by an organism resistant
CC to antibiotics that affect peptidoglycan synthesis. They may also be
CC used in screening assays, in vaccines, and in diagnostic assays. BUR
CC nucleotides are useful in gene therapy applications to modulate BUR
CC expression, and to detect BUR expression in a biological sample.
CC
XX
XX Sequence 481 BP; 139 A; 95 C; 100 G; 147 T; 0 other:
XX
XX Query Match 100.0%; Score 481; DB 24; Length 481;
XX Best Local Similarity 100.0%; Pred. No. 1,7e-129;
XX Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TGCCCTCATGTGAAGTATGATGCTATTGACACTATCCCTTTACCACGGCTCAACAGTTT 60
DB 1 TGCCTCTCATTTGAAGTATGATGCTATTGACACTATCCCTTTACCACGGCTCAACAGTTT 60
QY 61 AATAACCTGCCAGCAATTAAGGATGTTTAACTTAAGTCAAAAAAATAGCGAATTTTC 120
DB 61 AATAACCTGCCAGCAATTAAGGATGTTTAACTTAAGTCAAAAAAATAGCGAATTTTC 120
QY 121 CACGCAACAAAAGCTAAATATCGCAAAAACCTCAGTAAATCTTGCTGAGCTATTATT 180
DB 121 CACGCAACAAAAGCTAAATATCGCAAAAACCTCAGTAAATCTTGCTGAGCTATTATT 180
QY 181 GCTAAGTAACATTTTACCCCGTAAGTTAATGATCAATCAAGAGATGTGGCTGTAAT 240
DB 181 GCTAAGTAACATTTTACCCCGTAAGTTAATGATCAATCAAGAGATGTGGCTGTAAT 240
QY 241 GAATGCTTTAATGAATTAACAGTTGGATGCTTTCGTTTCAGTCAATCTTCTTGG 300
DB 241 GAATGCTTTAATGAATTAACAGTTGGATGCTTTCGTTTCAGTCAATCTTCTTGG 300
QY 301 CGTGGCAGTGCATTTGACCACTATCAGCCACCTGTAACAGAGTGGTGGTAAACACAA 360
DB 301 CGTGGCAGTGCATTTGACCACTATCAGCCACCTGTAACAGAGTGGTGGTAAACACAA 360
QY 361 GTTAACTGTGCATTTGTGAGCGACATGCTGCCGCTCGGGTGCATTTTTCACTTGTCG 420
DB 361 GTTAACTGTGCATTTGTGAGCGACATGCTGCCGCTCGGGTGCATTTTTCACTTGTCG 420
QY 421 GATATTAAACCGAATTTATTAATCTTGTCGATGTTGTGGTGAATGTCAGGATAGTA 480
DB 421 GATATTAAACCGAATTTATTAATCTTGTCGATGTTGTGGTGAATGTCAGGATAGTA 480
QY 481 A 481
DB 481 A 481
XX
XX RESULT 2
XX ABA90521/C
XX ID ABA90521 standard; DNA; 2365589 BP.
XX
XX AC ABA90521;
XX
XX DT 16-MAY-2002 (first entry)
XX
XX DE Genomic sequence of Lactococcus lactis IL1403.
XX
XX KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds-
XX Lactococcus lactis IL1403.
XX
XX OS FR2807446-A1.
XX
XX PN 12-OCT-2001.
XX
XX

```


XX	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW	acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW	gene; ds.
XX	
XX	Homo sapiens.
OS	
PN	W020020928-A2.
PD	
PD	03-JAN-2002.
XX	
XX	02-JUL-2001; 2001WO-EP07537.
PR	
PR	30-JUN-2000; 2000DE-1032529.
PR	01-SEP-2000; 2000DE-1043826.
XX	
PA	(EPIG-) EPIGENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K;
XX	
DR	WPI: 2002-130909/17.
PT	
PT	Nucleic acid comprising fragment of chemically modified gene, useful
PT	for diagnosis and treatment of diseases associated with abnormal
PT	cytosine methylation -
XX	
PS	Claim 1; SEQ ID NO 341; 32pp + Sequence Listing; German.
CC	
CC	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC	diseases. The present sequence is a gene of the invention.
XX	
XX	Sequence 9997 BP; 2290 A; 91 C; 2181 G; 5435 T; 0 other;
XX	
XX	Query Match 8.1%; Score 38.8; DB 24; Length 9997;
XX	Best Local Similarity 54.1%; Pred. NO. 0.77;
XX	Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY	13 AAGTATGATGGCTATTTCACACTATTCCTTACCACGCTCAACAGTTTATAACCTGCCA 72
DB	5923 AAGTATTATTAACCAATATACCAACAAACCAACACACTTAATCTTTACATATAAACATA 5864
QY	73 GCAATAGAGGATGTGGTTTAAGCTTAAGTCAAAAAAATAGCAATTTTCCACGACACAAA 132
DB	5863 TATATTTCACATATTTATATAAATTAACTCAAAAAACACACAAATTTCTCCACACACAAA 5804
QY	133 GCTAATATATCGCAAAAACCTCAGTAA 158
DB	5803 AAAAAATACATAAAAAAAAATAA 5778
XX	
XX	RESULT 5
XX	ABL34484/C
ID	ABL34484 standard; DNA; 9997 BP.
XX	
XX	ABL34484;
AC	
XX	
DT	26-MAR-2002 (first entry)
XX	
XX	Human metastasis associated gene SEQ ID NO: 37.
DE	
XX	Metastasis associated gene; cytosine; gene therapy; cancer;
KW	cytosine methylation; gene; ds.
KW	
XX	
OS	Homo sapiens.
XX	
XX	W0200177376-A2.
PN	
XX	

PD	18-OCT-2001.
XX	
PF	06-APR-2001; 2001WO-EP03970.
XX	
PR	06-APR-2000; 2000DE-1019058.
XX	
PR	07-APR-2000; 2000DE-1019173.
XX	
PR	30-JUN-2000; 2000DE-1032529.
XX	
PR	01-SEP-2000; 2000DE-1043826.
XX	
PA	(EPiG-) EPIGENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K;
XX	
DR	WPI; 2002-010922/01.
XX	
PT	New nucleic acid derived from chemically treated metastasis genes,
XX	
PT	useful for diagnosis of cancers by analysis of cytosine methylation,
XX	
PT	also for treatment -
XX	
PS	Claim 1; SEQ ID NO 37; 23pp + Sequence Listing; English.
XX	
CC	The present invention provides a number of human metastasis associated
XX	
CC	genes which are modified by cytosine methylation. The sequences can be
XX	
CC	used in the diagnosis and treatment of cancer. The present sequence is
XX	
CC	one of the genes of the invention.
XX	
SQ	Sequence 9997 BP; 2290 A; 91 C; 2181 G; 5435 T; 0 other;
XX	
Query Match	8.1%; Score 38.8; DB 24; Length 9997;
XX	
Best Local Similarity	54.1%; Pred. No. 0.77;
XX	
Matches	79; Conservative 0; Mismatches 67; Indels 0; Gaps 0
XX	
QY	13 AAGTATGATGGCTATTGTGACACTATTCCTTACCACGCTCAACAGTTTAACTGCGCA 72
XX	
DB	5923 AATATTATTATACCAACATTAACACAAACAAACACACCTTAATCTTTACAAATMAACATA 5864
XX	
QY	73 GCATTAAGGAGTGTCTTTAACTTAGTCAAAAAAATAGCAATTTTCCACGACAAAA 132
XX	
DB	5863 TATATTTCACATTTATTAATAATTAACCTCAAAAACACACATTTCTCCACACACAAAA 5804
XX	
QY	133 GCTAATATGCGCAAAAACCTCAGTAA 158
XX	
DB	5803 AAAAAATTAACATAAAAAAATAA 5778
XX	
RESULT 6	
XX	
ABK31511/C	
XX	
ID	ABK31511 standard; DNA; 47108 BP.
XX	
AC	ABK31511;
XX	
DT	23-APR-2002 (first entry)
XX	
DE	Signal transduction associated gene modified complementary DNA #177.
XX	
XX	Human; signal transduction associated gene; cytosine methylation state;
XX	CPG island; signal transduction associated disease; solid tumour; cancer;
XX	antitumour; cytostatic; mutant; ds.
XX	
OS	Homo sapiens.
XX	
OS	Synthetic.
XX	
PN	WO200200926-A2.
XX	
PD	03-JAN-2002.
XX	
PF	29-JUN-2001; 2001WO-EP07472.
XX	
PR	30-JUN-2000; 2000DE-1032529.
XX	
PR	01-SEP-2000; 2000DE-1043826.
XX	
PA	(EPiG-) EPIGENOMICS AG.
XX	

PI Olek A, Piepenbrock C, Berlin K;
XX WPI: 2002-147896/19.
XX
XX Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction -
XX
XX
XX Claim 1; SEQ ID NO 354; 24pp; English.
XX
XX The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or
CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers
CC for detecting the cytosine methylation state (CpG Islands) of these
CC genes, and a method for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with signal transduction.
CC The genomic DNA can be obtained from cells or cellular components which
CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
CC histologic object slides, and all their possible combinations. The
CC sequences of the invention are useful for the diagnosis and therapy of
CC diseases associated with signal transduction e.g. solid tumours and
CC cancer. ABK3158-ABK31545 represent chemically pretreated genomic DNA
CC sequences of different genes associated with signal transduction, or
CC their complementary sequences.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
XX Sequence 47108 BP: 15349 A; 429 C; 8736 G; 22594 T; 0 other;
SQ
Query Match 8.0%; Score 38.4; DB 24; Length 47108;
Best Local Similarity 48.2%; Pred. No. 1.8;
Matches 108; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
OY 43 ACCACGCTCAACGATTATTAACCTGCCAGCAATAGGATGCTTTAACTTAACCTCA 102
DB 19108 ACACCCACCAACCTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 19049
OY 103 AAAAAATAGCGAATTTTCCACAGCAACAAAGCTAATATGCAAAAACTCAGTAAAAAT 162
DB 19048 TAAAAACAAAAAATAAAAAATCCCAAAATTAACAAATATTAATTTCCCTAAAACT 18989
OY 163 CTGCTGAGCTATTAATGCTAAGTAACATTACCCCTGAAGTTAATGATCAATCAAG 222
DB 18988 AATATATTAATTAATTAACCTTAATCAACAATTAATTAATTAATTTTAATTA 18929
OY 223 AGAGATGTGGGCTGATGAATGCTTATTAATTAACGATT 266
DB 18928 ACCCAATTAACCACTTAATAATCCATTAATAAATTAATTAACCTTT 18885

RESULT 7
ABL70143/C
ID ABL70143 standard; DNA: 15881 BP.
XX
XX ABL70143;
XX
XX 01-JUL-2002 (first entry)
XX
XX Chemically treated cell signalling DNA sequence#17.
XX
XX Cell signalling: cytosine methylation; cell signalling disease;
KW cancer; tumour; cytosine; ds.
XX
XX unidentified.
XX
XX WO200202807-A2.
XX
XX 10-JAN-2002.
XX
XX

PF 29-JUN-2001; 2001WO-EP07471.
XX
XX 30-JUN-2000; 2000DE-1032529.
PR
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI: 2002-154758/20.
XX
XX
XX Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signalling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signalling -
XX
XX
XX Claim 1; SEQ ID NO 33; 24pp+sequence listing; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX
XX Sequence 15881 BP: 4515 A; 204 C; 3588 G; 7574 T; 0 other;
SQ
Query Match 7.9%; Score 37.8; DB 24; Length 15881;
Best Local Similarity 52.2%; Pred. No. 1.8;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
OY 102 AAAAAATAGCGAATTTTCCACAGCAACAAAGCTAATATGCAAAAACTCAGTAAAAA 161
DB 14565 AAAAAACAAATCAAAAAACACAAAAAAGACATTAATTAATTAATTAATTAATTAATTA 14506
OY 162 TCTTGCTGAGCTATTAATGCTAAGTAACATTACCCCTGAAGTTAATGATCAATCAA 221
DB 14505 CCTCCATTCATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 14446
OY 222 GAGGATGTGGGCTGATGAATGCTTATTAATTAACGATT 262
DB 14445 AATATATCTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 14405

RESULT 8
ABL32260/C
ID ABL32260 standard; DNA: 15881 BP.
XX
XX ABL32260;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 233.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; neoplastic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antineumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
XX Homo sapiens.
OS

XX W0200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07537.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
XX 01-SEP-2000; 2000DE-1043626.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
XX PT for diagnosis and treatment of diseases associated with abnormal
XX PT cytosine methylation -
XX
XX Claim 1; SEQ ID NO 233; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX CC genes which are modified by the methylation of cytosines. The sequences
XX CC can be used in the diagnosis and treatment of immune system disorders,
XX CC including eye diseases such as retinopathy, neovascular glaucoma and
XX CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 15881 BP; 4515 A; 204 C; 3588 G; 7574 T; 0 other;

Query Match 7.9%; Score 37.8; DB 24; Length 15881;
Best Local Similarity 52.2%; Pred. No. 1.8;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 102 AAAAAAATGCGAATTTCCAGACAAAGCTAAATATCGAAAAACCTCAGTAAAA 161
DB 14565 AAAAAAATGCGAATTTCCAGACAAAGCTAAATATCGAAAAACCTCAGTAAAA 14506
QY 162 TCTTGGTGGAGCTATTATGCTAAGTAACATTTACCCCTGAAGTTAATGATCAATCA 221
DB 14505 CCTCCATTCGAATTTAATTAATAATAAATTTTCCTCTAATAAAAAAATAAATAA 14446
QY 222 GAGAGATGCGGCTGTAATGAATCGTCTTAATTAATTAACA 262
DB 14445 AATATATCTATTTTAAATAAACCTCTAATCTTTAATA 14405

RESULT 9
AAS61069/c
ID AAS61069 standard; DNA; 15881 BP.
XX
XX AAS61069;
XX
XX 29-JAN-2002 (first entry)
XX
XX Human gene regulation-associated gene oligonucleotide #24.
XX
XX Human: Gene regulation-associated gene; severe combined immunodeficiency;
XX KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
XX KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
XX KW renal disease; Preeclampsia; cardiac allograft vascular disease;
XX KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
XX KW immunostimulant; cardiac; anti-inflammatory; coagulant; antiasthmatic;
XX KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cystostatic.
XX
XX Homo sapiens.
XX
XX W020017375-A2.
XX
XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-EP03968.
XX
XX 06-APR-2000; 2000DE-1019058.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
XX 01-SEP-2000; 2000DE-1043626.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-017470/02.
XX
XX New nucleic acid sequences from chemically modified genes associated
XX PT with gene regulation, useful for analysing cytosine methylations for
XX PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
XX PT disease -
XX
XX Claim 1; SEQ ID NO 25; 26pp; English.
XX
XX The invention relates to 224 nucleic acid sequences comprising at least
XX CC 18 bases of a chemically pretreated gene associated with gene regulation
XX CC selected from 43 known genes (or complementary sequences). The
XX CC chemical pretreatment converts cytosine bases unmethylated at the
XX CC 5-position to uracil or another base with hybridisation behaviour
XX CC dissimilar to cytosine, to enable analysis of cytosine methylations.
XX CC The DNA sequences, oligomers (or sets/arrays) and method are
XX CC useful in the diagnosis of diseases (or predisposition to diseases)
XX CC associated with gene regulation and in therapy of such diseases, by
XX CC enabling analysis of the cytosine methylation patterns of such genes,
XX CC kits are provided. They are especially useful in diagnosis
XX CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
XX CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
XX CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
XX CC preclampsia, graft versus-host disease. The present sequence is a
XX CC sequence included in the sequence data for this specification and is
XX CC associated with the human gene regulation-associated genes.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 15881 BP; 4515 A; 204 C; 3588 G; 7574 T; 0 other;

Query Match 7.9%; Score 37.8; DB 24; Length 15881;
Best Local Similarity 52.2%; Pred. No. 1.8;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 102 AAAAAAATGCGAATTTCCAGACAAAGCTAAATATCGAAAAACCTCAGTAAAA 161
DB 14565 AAAAAAATGCGAATTTCCAGACAAAGCTAAATATCGAAAAACCTCAGTAAAA 14506
QY 162 TCTTGGTGGAGCTATTATGCTAAGTAACATTTACCCCTGAAGTTAATGATCAATCA 221
DB 14505 CCTCCATTCGAATTTAATTAATAATAAATTTTCCTCTAATAAAAAAATAAATAA 14446
QY 222 GAGAGATGCGGCTGTAATGAATCGTCTTAATTAATTAACA 262
DB 14445 AATATATCTATTTTAAATAAACCTCTAATCTTTAATA 14405

RESULT 10
ABK31182/c
ID ABK31182 standard; DNA; 15881 BP.
XX
XX ABK31182;
XX
XX 23-APR-2002 (first entry)
XX
XX Signal transduction associated gene modified DNA #13.
XX
XX Human; signal transduction associated gene; cytosine methylation state;
XX
XX

KM CPG island: signal transduction associated disease; solid tumour; cancer;
KM antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
XX Synthetic.
XX
PN WO200200926-A2.
XX
PD 03-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP07472.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Plepenbrock C, Berlin K;
XX
DR WPI: 2002-147896/19.
XX
PT Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction -
XX
PS Claim 1; SEQ ID No 25; 24pp; English.
XX
CC The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or
CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers
CC for detecting the cytosine methylation state (CPG islands) of these
CC genes, and a method for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with signal transduction.
CC The genomic DNA can be obtained from cells or cellular components which
CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
CC histologic object slides, and all their possible combinations. The
CC sequences of the invention are useful for the diagnosis and therapy of
CC diseases associated with signal transduction e.g. solid tumours and
CC cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA
CC sequences of different genes associated with signal transduction, or
CC their complementary sequences.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 15881 BP; 4515 A; 204 C; 3588 G; 7574 T; 0 other;
XX
Query Match 7.9%; Score 37.8; DB 24; Length 15881;
Best Local Similarity 52.2%; Pred. No. 1.8;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
XX
QY 102 AAAAAATAGCGAATTTCCACAGCAAAAGCTAATATCGAATAAAGCTCAGTAAAA 161
DB 14565 AAAAAATAGCGAATTTCCACAGCAAAAGCTAATATCGAATAAAGCTCAGTAAAA 14506
QY 162 TCTTGGTGGAGCTATATTGCTAAGTACATTTACCCCTGAAGTTAATGATCAATCAA 221
DB 14505 CCGCCATTCGAATTTAATTAATATAAAATTTTCTCTAATTAATAAAATATAA 14446
QY 222 GAGAGATGTGGGCTGAATGAATCGTTAATGAATTAACA 262
DB 14445 AATATATCTATTTTAAATAAAGCTCTTACTTTAATA 14405
XX
RESULT 11
AAC85009
ID AAC85009 standard; DNA; 15456 BP.
XX
AC AAC85009;
XX

DT 08-MAY-2001 (first entry)
XX
XX Nucleotide sequence of bovine PIV3 SF strain.
DE
XX
XX Parainfluenza virus; PIV; infectious; human; nucleocapsid protein; BPIV;
KM nucleocapsid phosphoprotein; polymerase protein; human; bovine; HPIV;
KM chimeric; vaccine; immune response; HPIV1; HPIV2; HPIV3; BPIV3; ds.
XX
OS Bovine parainfluenza virus.
XX
XX WO200104320-A1.
XX
XX 18-JAN-2001.
XX
XX 16-JUN-2000; 2000WO-US17066.
XX
XX 09-JUL-1999; 99US-0143134.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Schmidt AC, Skiadopoulos MH, Collins PL, Murphy BR, Bailly JE;
PI Durbin AP;
XX
XX WPI: 2001-081053/09.
XX
XX Isolated human-bovine chimeric parainfluenza virus (PIV), useful in an
XX attenuated vaccine to elicits an immune response against one or more
XX virus(es) selected from human PIV1 (HPIV1), HPIV2 and HPIV3 -
XX
XX Disclosure: Fig 7A-G; 148pp; English.
XX
CC The invention relates to an isolated human-bovine chimeric parainfluenza
CC virus (PIV) that is infectious and attenuated in humans. The virus
CC comprises a major nucleocapsid protein, a nucleocapsid phosphoprotein,
CC a large polymerase protein, and a partial or complete PIV background
CC genome, or antigenome of a human PIV (HPIV) or bovine PIV (BPIV),
CC combined with one or more heterologous gene(s) or genome segment(s) of a
CC different PIV to form a human-bovine chimeric PIV genome or antigenome.
CC The chimeric PIV is useful in attenuated vaccines to elicit immune
CC response against one or more virus(es) selected from HPIV1, HPIV2 and
CC HPIV3 and most preferably against HPIV3. The present sequence represents
CC the nucleotide sequence of bovine PIV3 SF strain.
XX
SQ Sequence 15456 BP; 5872 A; 2656 C; 2935 G; 3993 T; 0 other;
XX
Query Match 7.7%; Score 37.2; DB 22; Length 15456;
Best Local Similarity 54.3%; Pred. No. 2.6;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
XX
QY 208 AATGATCAATCAAGAGATGTGGCTGATGAATCGTTATTGAATTAACAGGTTG 267
DB 3926 AATGGAAGGTCAAAAAGACAGATGAGCATTAAGTGAATCAAGATGATCAATGA 3985
QY 268 GATCGTTTGTGCGTTTCAGTCAATCTTCTGCGGAGGAGACATTAAGCAACATCA 327
DB 3986 CAAAGTTTGTGCTGTGATCATTTGCCACTTGCGTTGCTAGTATACATGGAATGATCA 4045
QY 328 GCCACCTGAACAGAGTGC 345
DB 4046 GGAACCTCAACAGGCTGC 4063
XX
RESULT 12
AAD27766
ID AAD27766 standard; DNA; 15456 BP.
XX
AC AAD27766;
XX
XX 18-APR-2002 (first entry)
DT
XX Bovine parainfluenza virus 3 (BPIV3) SF DNA.
DE
XX Parainfluenza virus; PIV; nucleocapsid protein; polymerase protein;
XX

KW HIV; human PIV, nucleocapsid phosphoprotein; bovine PIV3; BPIV3;
 KW prophylaxis; infection; vaccine; immune response; immunostimulant;
 KW virucide; SF DNA; ds.
 XX
 OS Bovine parainfluenza virus.
 XX
 PN WO200202605-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 05-JUL-2001; 2001WO-US21527.
 XX
 PR 05-JUL-2000; 2000US-215809P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Skladopoulos MH, Collins PL, Murphy BR, Schmidt AC;
 XX
 DR WPI: 2002-140084/18.
 XX
 PT Isolated human-bovine chimeric parainfluenza virus (PIV), useful in an
 PT attenuated vaccine to elicits an immune response against one or more
 PT virus(es) selected from human PIV1 (HPV1), HPV2 and HPV3 -
 XX
 PS Example 1; Fig 2; 154pp; English.
 XX
 CC The present invention relates to chimeric human-bovine parainfluenza
 CC viruses (PIV) that are infectious and attenuated in humans and other
 CC mammals. The viruses comprise a major nucleocapsid protein (N), a
 CC nucleocapsid phosphoprotein (P), a large polymerase protein (L) and
 CC a partial or complete PIV background genome or antigenome of a human
 CC PIV (HPV) or bovine PIV (BPIV), combined with one or more heterologous
 CC genes or genome segments of a N, P, L, or M gene of a different PIV to
 CC form a human-bovine chimeric PIV genome or antigenome. The invention
 CC also relates to methods and compositions incorporating human-bovine
 CC chimeric PIV for prophylaxis and treatment of PIV infection. The
 CC chimeric PIVs are useful in attenuated vaccines to elicit an immune
 CC response against one or more viruses selected from HPV1, HPV2 and
 CC HPV3. The present sequence is bovine parainfluenza virus 3 (BPIV3)
 CC SF DNA. This sequence is used in the exemplification of the invention.
 XX
 SQ Sequence 15456 BP; 5871 A; 2653 C; 2936 G; 3996 T; 0 other;
 XX
 Query Match 7.7%; Score 37.2; DB 24; Length 15456;
 Best Local Similarity 54.3%; Pred. No. 2.6;
 Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
 XX
 QY 208 AATGATCATCAAGAGAGATGGGCTGAATGATCGTCTTATGATTAACAGGTG 267
 DB 3926 AATGGAAGGTCAAAAGACAGGTATGAGCATTAAGTATGATGATCAAGTTA 3985
 QY 268 GATCGTCTTCTGCTTCAGTCAATCTTGTGGCGGAGTACATTGACAACTATCA 327
 DB 3986 CAAGGTTTGGCTCTGATCATTTGGCAGTGGCTGAGTATACAGGAATATCA 4045
 QY 328 GCCACTGAACAGAGTGC 345
 DB 4046 GGAACCTCTACAGGCTGC 4063
 XX
 RESULT 13
 ID ABL34168/c
 AC ABL34168 standard; DNA; 5914 BP.
 XX
 AC ABL34168;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 2141.
 XX
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antiamebic; cytosinetic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW

KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI: 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 2141; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 5914 BP; 1235 A; 162 C; 1616 G; 2900 T; 1 other;
 XX
 Query Match 7.7%; Score 36.8; DB 24; Length 5914;
 Best Local Similarity 51.2%; Pred. No. 2.4;
 Matches 86; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
 XX
 QY 31 ACATGATCCTTTATCCACGCTTCACAGTTTAATTAACCTGCCAATTAAGGATGTTT 90
 DB 3073 AAACTCCCATTCACACAGCTCAATCAATTTAAACCAATTAACCTTTATCCATCAATC 3014
 QY 91 TAACTTAAGTCAAAAAAATAGCAATTTTCCAAAGCAAAAGCTAATATCGCAAAAC 150
 DB 3013 TAACTTCTAATAATTAATTTCTTAACAACCCACGAAATTAATATCCGCTCGAAAAAT 2954
 QY 151 CTCAGTAAATATCTGCTGAGCTATTTATGCTAAGTAACTTTATCC 198
 DB 2953 TCCAAACCAAAATATCAAAAGCTATTTACATCTAATAATACGATATCCC 2906
 XX
 RESULT 14
 ID ABL70321/c
 AC ABL70321 standard; DNA; 6310 BP.
 XX
 AC ABL70321;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Chemically treated cell signalling DNA sequence#106.
 XX
 KW Cell signalling; cytosine methylation; cell signalling disease;
 KW cancer; tumour; cytosinetic; ds.
 XX
 OS Unidentified.
 XX
 PN WO200202807-A2.
 XX

XX 10-JAN-2002.
PD
XX
XX 29-JUN-2001; 2001WO-EP07471.
PF
XX
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-154758/20.
DR
XX
XX Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signalling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signalling -
XX
XX
PS Claim 1; SEQ ID NO 211; 24pp+sequence listing; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records AB170111-AB170626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX
XX
SQ Sequence 6310 BP; 1426 A; 213 C; 1583 G; 3088 T; 0 other;
XX
XX
XX Query Match 7.7%; Score 36.8; DB 24; Length 6310;
XX Best local Similarity 54.4%; Pred. No. 2.5;
XX Matches 74; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
XX
XX 92 AACTTAAGTCAAAAAATGCGCAATTTTCCACGACAAAGCTAATATCGCAAAACC 151
DB 2449 AATTAACCAAAAAAATATCGAAAAATTTCCAAAAAACAATAATACAAAAACC 2390
XX
XX 152 TCAGTAAATCTTGCTGGAGTATATTCCTAAGTAACATTACCCCGGAAGTTAAG 211
DB 2389 TAAAAAATATCAAACTTAATATTTAAAAAATAACGAAAAAACCCTCAATTAATTAAATA 2330
XX
XX 212 GATCAATCAAGAGAGA 227
DB 2329 AACAAAAAATAAAAAA 2314
XX
XX
XX RESULT 15
XX AAS61268/C
XX ID AAS61268 standard; DNA; 6310 BP.
XX
XX AAS61268;
XX
XX 29-JAN-2002 (first entry)
XX
XX Human gene regulation-associated gene oligonucleotide #223.
XX
XX Human; Gene regulation-associated gene; severe combined immunodeficiency;
KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
KW renal disease; Preeclampsia; cardiac allograft vascular disease;
KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
KW immunostimulant; cardiac; antiinflammatory; coagulant; antiasthmatic;
KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.

XX Homo sapiens.
OS
XX
XX WO200177375-A2.
PN
XX
XX 18-OCT-2001.
PD
XX
XX 06-APR-2001; 2001WO-EP03968.
PF
XX
XX 06-APR-2000; 2000DE-1019058.
PR
XX 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-017470/02.
DR
XX
XX New nucleic acid sequences from chemically modified genes associated
PT with gene regulation, useful for analysing cytosine methylations for
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
PT disease -
XX
XX
XX Disclosure; SEQ ID No 229; 26pp; English.
XX
XX The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The
CC chemical pretreatment converts cytosine bases unmethylated at the
CC 5-position to uracil or another base with hybridisation behaviour
CC dissimilar to cytosine, to enable analysis of cytosine methylations.
CC The DNA sequences, oligomers (or sets/arrays) and method are
CC useful in the diagnosis of diseases (or predisposition to diseases)
CC associated with gene regulation and in therapy of such diseases, by
CC enabling analysis of the cytosine methylation patterns of such genes,
CC kits are provided. They are especially useful in diagnosis
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC preeclampsia, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pt_sequences
XX
XX
SQ Sequence 6310 BP; 1426 A; 213 C; 1583 G; 3088 T; 0 other;
XX
XX
XX Query Match 7.7%; Score 36.8; DB 24; Length 6310;
XX Best local Similarity 54.4%; Pred. No. 2.5;
XX Matches 74; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
XX
XX 92 AACTTAAGTCAAAAAATGCGCAATTTTCCACGACAAAGCTAATATCGCAAAACC 151
DB 2449 AATTAACCAAAAAAATATCGAAAAATTTCCAAAAAACAATAATACAAAAACC 2390
XX
XX 152 TCAGTAAATCTTGCTGGAGTATATTCCTAAGTAACATTACCCCGGAAGTTAAG 211
DB 2389 TAAAAAATATCAAACTTAATATTTAAAAAATAACGAAAAAACCCTCAATTAATTAAATA 2330
XX
XX 212 GATCAATCAAGAGAGA 227
DB 2329 AACAAAAAATAAAAAA 2314
XX
XX
XX Search completed: February 7, 2003, 19:59:38
XX Job time : 1034 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 19:24:14 ; Search time 86 Seconds

(without alignments)
1715.250 Million cell updates/sec

Title: US-09-828-456-1

Perfect score: 481
Sequence: 1 tgcctcatgaagatga.....gtgatgcaggaagataa 481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents.NA: *
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCITUS.COMB.seq: *
6: /cgn2_6/ptodata/1/lna/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35.6	7.4	443	4	US-09-397-787-272
2	33.6	7.0	4078	2	US-08-960-022-3
3	32.4	6.7	7568	2	US-08-694-869-2
4	32.4	6.7	7568	3	US-09-349-546-2
5	32	6.7	2861	1	US-08-299-953-1
6	32	6.7	2861	1	US-08-459-415-1
7	32	6.7	2861	4	US-09-066-687-1
8	32	6.7	2861	5	PCIT-US95-11231-1
9	32	6.7	3881	1	US-08-299-953-2
10	32	6.7	3881	1	US-08-459-415-2
11	32	6.7	3881	4	US-09-066-687-2
12	32	6.7	3881	5	PCIT-US95-11231-2
13	31.8	6.6	536	1	US-08-941-568-1
14	31.8	6.6	536	2	US-08-911-020-1
15	31.8	6.6	1449	3	US-09-135-639-1
16	31.6	6.6	1896	4	US-09-149-476-55
17	31.6	6.6	12886	4	US-09-453-702B-14
18	31.4	6.5	9048	3	US-08-973-273-4
19	31.2	6.5	4220	1	US-08-832-883-66
20	31.2	6.5	4220	2	US-08-832-877-66
21	30.6	6.4	1782	4	US-09-069-023-37
22	30.6	6.4	658	4	US-08-998-416-595
23	30.6	6.4	1102	4	US-08-821-994-86
24	30.6	6.4	1434	4	US-08-821-994-86
25	30.6	6.4	1717	4	US-09-149-476-124
26	30.6	6.4	2274	4	US-09-388-743-17
27	30.6	6.4	5232	3	US-09-212-971-3

c 28	30.6	6.4	5232	3	US-08-800-929A-3	Sequence 3, App11
c 29	30.6	6.4	5232	4	US-09-617-053A-3	Sequence 3, App11
c 30	30.4	6.3	2886	2	US-08-687-080-55	Sequence 55, App1
c 31	30.4	6.3	9048	3	US-08-973-273-4	Sequence 4, App11
c 32	30.2	6.3	216	4	US-09-134-001C-1302	Sequence 1302, Ap
c 33	30.2	6.3	1903	1	US-07-961-522-3	Sequence 3, App11
c 34	30.2	6.3	1903	1	US-08-217-438-3	Sequence 3, App11
c 35	30.2	6.3	1903	1	US-08-321-978-3	Sequence 3, App11
c 36	30.2	6.3	1903	2	US-08-710-584-3	Sequence 3, App11
c 37	30.2	6.3	6152	4	US-08-973-462-1	Sequence 1, App11
c 38	30	6.2	271	2	US-08-731-272A-29	Sequence 29, App1
c 39	30	6.2	1287	4	US-09-134-001C-2222	Sequence 2222, Ap
c 40	30	6.2	2035	3	US-09-022-699-3	Sequence 3, App11
c 41	30	6.2	2968	3	US-09-022-699-1	Sequence 1, App11
c 42	30	6.2	4285	4	US-09-410-464-1	Sequence 1, App11
c 43	29.8	6.2	1930	2	US-08-455-968E-4	Sequence 4, App11
c 44	29.8	6.2	2033	2	US-08-455-968E-9	Sequence 9, App11
c 45	29.6	6.2	3719	1	US-08-920-812-10	Sequence 10, App1

ALIGNMENTS

```
RESULT 1
US-09-397-787-272
Sequence 272, Application US/09397787
Patent No. 6468758
GENERAL INFORMATION:
APPLICANT: Benson, Darin R.
APPLICANT: Lodes, Michael J.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
FILE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.466C2
CURRENT APPLICATION NUMBER: US/09/397,787
NUMBER OF SEQ ID NOS: 334
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 272
LENGTH: 443
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(443)
OTHER INFORMATION: n = A,T,C or G
US-09-397-787-272

Query Match
Best Local Similarity 7.4%; Score 35.6; DB 4; Length 443;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 100 TCAAAAAAATGAGCAATTTTCCACGACCAAAAGCTTAATGCAAAACCTCAGTAA 159
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 207 TCAACCAAAATGAGCAATTTTCCACGACCAAAAGCTTAATGCAAAACCTCAGTAA 266
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 160 AATCTTGCTGGAGCAATTTTCCACGACCAAAAGCTTAATGCAAAACCTCAGTAA 219
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 267 AATGCTTGCTGGAGCAATTTTCCACGACCAAAAGCTTAATGCAAAACCTCAGTAA 326
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 220 AAGAGATGTGGGCTCT 237
Db 327 TAAATCATTTGAGGGCT 344

RESULT 2
US-08-960-022-3/C
Sequence 3, Application US/08960022
Patent No. 5976837
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
```

```

; APPLICANT: Lavallee, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,022
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4078 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-960-022-3

Query Match
Best Local Similarity 7.0%; Score 33.6; DB 2; Length 4078;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 53 AACAGTTTAATACCGCAGCAATAGGAGTGTGTTAACTTAAGTCAAAAAAATAGC 112
DB 3720 AAAAGAGTCTTATCTCTCCCAAAATAGAAATTTGTATCTATAGTCAAAACAATAA 3661
QY 113 GAATTTCCACGACGACAAAGCTAATATCGCAAAA 148
DB 3660 CATTTTGAATTAATAAAAGTGAAGAACTTGAAA 3625

RESULT 3
US-08-694-869-2
; Sequence 2, Application US/08694869
; Patent No. 5994123
; GENERAL INFORMATION:
; APPLICANT: Olszewski, N.
; APPLICANT: Tzafir, I.
; APPLICANT: Somers, D. A.
; APPLICANT: Lockhart, B.
; APPLICANT: Torbert, K.
; TITLE OF INVENTION: SUGARCANE BACILLIFORM VIRUS
; TITLE OF INVENTION: PROMOTER
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kiuth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,869
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.369US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-694-869-2

Query Match
Best Local Similarity 6.7%; Score 32.4; DB 2; Length 7568;
Matches 63; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 53 AACAGTTTAATACCGCAGCAATAGGAGTGTGTTAACTTAAGTCAAAAAAATAGC 112
DB 5962 AACATATTTATTGAAGAGTGGAGATTGTGACGAGAGTGAAGACAAAGAGGTC 6021
QY 113 GAATTTCCACGACGACAAAGCTAATATCGCAAAAACCTCAATAAATCTTG 166
DB 6022 TTAATCTGCGTAGCAACACTGAACTATGCCAGAAACCAATCAAGATATAG 6075

RESULT 4
US-09-349-546-2
; Sequence 2, Application US/09349546
; Patent No. 6093569
; GENERAL INFORMATION:
; APPLICANT: Olszewski, N.
; APPLICANT: Tzafir, I.
; APPLICANT: Somers, D.A.
; APPLICANT: Lockhart, B.
; APPLICANT: Torbert, K.
; TITLE OF INVENTION: Sugarcane bacilliform virus promoter
; FILE REFERENCE: 600.369US2
; CURRENT APPLICATION NUMBER: US/09/349,546
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: US 08/694,869
; EARLIER FILING DATE: 1996-08-09
; EARLIER APPLICATION NUMBER: PCT/IB97/01338
; EARLIER FILING DATE: 1997-08-13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 7568
; TYPE: DNA
; ORGANISM: sugarcane bacilliform virus
; US-09-349-546-2

Query Match
Best Local Similarity 6.7%; Score 32.4; DB 3; Length 7568;
Matches 63; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 53 AACAGTTTAATACCGCAGCAATAGGAGTGTGTTAACTTAAGTCAAAAAAATAGC 112
```

Db 5962 ACATATTTATTAAGAGATTGCAAGATTTCATACGAGAAAGTGAAGACAAAAGAGCTC 6021
QY 113 GAATTTTCCACAGACAAAAGCTAAATATTCGCAAAAACCTCAGTAAATCTTG 166
Db 6022 TTAATTCCTGGTAGCAACACTGACTATGCAGAAACACATCAAGATATG 6075

RESULT 5

US-08-299-953-1
; Sequence 1, Application US/08299953
; Patent No. 5646333
; GENERAL INFORMATION:
; APPLICANT: Dobres, Michael S. and Mandact, Sevnuur
; TITLE OF INVENTION: A Plant Promoter Useful for Directing the
; NUMBER OF INVENTION: Expression of Foreign Proteins to the Plant Epidermis
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5646333r1s
; STREET: One Liberty Place 46th. Floor
; City: Philadelphia
; STATE: PA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,953
; FILING DATE: Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: NOVA-0003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-564-8960
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2861 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-299-953-1

Query Match 6.7%; Score 32; DB 1; Length 2861;
Best Local Similarity 50.7%; Pred. No. 2.4;
Matches 77; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 89 TTTAACTTAAGTCAAAAAATAGCAATTTTCCACAGACAAAAGCTAAATATCGCAAAA 148
Db 1089 TATATCTAATATTAATTAATCAATTTTCAATGTTAAAGAAATTAATATGTTTATATAT 1148
QY 149 ACCTCAGTAAATAATCTTGCTGAGACTATTTATTCGTAAGTAACATTTACCCCTGAAGTTA 208
Db 1149 AACTCTGAATTTTATATATATCTTTTATTAATTAATTAATTTATCTTCGCATTTT 1208
QY 209 ATGATCAATCAAGAGAGATGGGCTGTAT 240
Db 1209 ATTAATATATAAGATATTTTGTATATAT 1240

RESULT 6
US-08-459-415-1
; Sequence 1, Application US/08459415
; Patent No. 5744334
; GENERAL INFORMATION:
; APPLICANT: Dobres, Michael S. and Mandact, Sevnuur

; TITLE OF INVENTION: A Plant Promoter Useful for Directing the
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5744334r1s
; STREET: One Liberty Place 46th. Floor
; City: Philadelphia
; STATE: PA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,415
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/299,953
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: NOVA-0003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-564-8960
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2861 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-459-415-1

Query Match 6.7%; Score 32; DB 1; Length 2861;
Best Local Similarity 50.7%; Pred. No. 2.4;
Matches 77; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 89 TTTAACTTAAGTCAAAAAATAGCAATTTTCCACAGACAAAAGCTAAATATCGCAAAA 148
Db 1089 TATATCTAATATTAATTAATCAATTTTCAATGTTAAAGAAATTAATATGTTTATATAT 1148
QY 149 ACCTCAGTAAATAATCTTGCTGAGACTATTTATTCGTAAGTAACATTTACCCCTGAAGTTA 208
Db 1149 AACTCTGAATTTTATATATATCTTTTATTAATTAATTAATTTATCTTCGCATTTT 1208
QY 209 ATGATCAATCAAGAGAGATGGGCTGTAT 240
Db 1209 ATTAATATATAAGATATTTTGTATATAT 1240

RESULT 7
US-09-066-687-1
; Sequence 1, Application US/09066687
; Patent No. 6339185
; GENERAL INFORMATION:
; APPLICANT: Dobres, Michael S. and Mandact, Sevnuur
; TITLE OF INVENTION: A Plant Promoter Useful for Directing the
; NUMBER OF INVENTION: Expression of Foreign Proteins to the Plant Epidermis
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6339185r1s
; STREET: One Liberty Place 46th. Floor
; City: Philadelphia
; STATE: PA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
STREET: One Liberty Place 46th. Floor
City: Philadelphia
STATE: PA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11231
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/299,953
FILING DATE: September 2, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: NOVA-0016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-564-8960
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3881 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US95-11231-2

Query Match 6.7%; Score 32; DB 5; Length 3881;
Best Local Similarity 50.7%; Pred. No. 2.8;
Matches 77; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 89 TTTACTTAAGTCAAAAATAGCAATTTCCAAAGCAAAAGCTTAATATCGCAAAA 148
DB 1089 TATATCATATATATATATATATTTTACAAATGTTAAAGATTAATATAGTTTATATAT 1148
QY 149 ACCTCAGTAAATCTTGCTGAGCTATATTGCTAAGTAACTTACCCCTGAAGTTA 208
DB 1149 AACTCTAAGATTTTATTTATTTTATTTTAAATATATATTAATTTATCTCCATTTT 1208
QY 209 ATGATCATATCAGAGAGATGTGGCTGTAT 240
DB 1209 ATTATATATTAAGATATTTTTCATATTAAT 1240

RESULT 13
US-08-341-568-1
Sequence 1, Application US/08341568
Patent No. 5661021
GENERAL INFORMATION:
APPLICANT: Buchert, Johanna
APPLICANT: Silka-aho, Matti
APPLICANT: Viikari, Liisa
APPLICANT: Penttila, Merja
APPLICANT: Saloheimo, Anu
APPLICANT: Marjatta, Rannu
TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching lig
TITLE OF INVENTION: pulps
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch and Birch
STREET: PO Box 747
CITY: Falls Church
STATE: VA
COUNTRY: US

ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/341,568
FILING DATE: 22-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr, Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 365-262P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 536 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Trichoderma reesei
STRAIN: QM9414
US-08-341-568-1

Query Match 6.6%; Score 31.8; DB 1; Length 536;
Best Local Similarity 50.3%; Pred. No. 1.3;
Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 18 TGATGCTATTTGACACTATCTTACCCAGCTCAACAGTTAATTAATCTGCACCAAT 77
DB 382 TGATCTCTGACATCTCATTCATGATTAACCTTCTTAATGATGCTAGTAAAC 441
QY 78 AAGGATGTTGTTTAACTTAAGTCAAAAATAGCAATTTTCCAAAGCAAAAGCTAA 137
DB 442 AAAGTTGTAGGAGTAAAGTAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAA 501
QY 138 ATATCGCAAAAGCTCACTAAATCTTGCTGGAG 172
DB 502 AAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCA 536

RESULT 14
US-08-911-020-1
Sequence 1, Application US/08911020
Patent No. 5854047
GENERAL INFORMATION:
APPLICANT: Buchert, Johanna
APPLICANT: Silka-aho, Matti
APPLICANT: Viikari, Liisa
APPLICANT: Penttila, Merja
APPLICANT: Saloheimo, Anu
APPLICANT: Marjatta, Rannu
TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching
TITLE OF INVENTION: pulps
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch and Birch
STREET: PO Box 747
CITY: Falls Church
STATE: VA
COUNTRY: US
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 19:42:18 ; Search time 98 Seconds

(without alignments)
2326.484 Million cell updates/sec

Title: US-09-828-456-1

Perfect score: 481
Sequence: 1 tgccttcattgaagatga.....gtgatgtcagagatagtaa 481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/us07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCr_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCrUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	481	100.0	481	10	US-09-828-456-1
2	38.6	8.0	1947	9	US-09-938-842A-3476
3	36.8	7.7	332	10	US-09-983-965-5260
4	36.8	7.7	7053	10	US-09-764-864-7292
5	35.6	7.4	443	10	US-09-876-889-272
6	35	7.3	66109	10	US-09-880-107-3768
7	35	7.3	198285	10	US-09-880-107-3814
8	34.2	7.1	1512	9	US-09-938-842A-3316
9	34.2	7.1	2000	10	US-09-887-576-859
10	34.2	7.1	2075	9	US-10-088-841-1110
11	33.6	7.0	4083	10	US-09-925-302-177
12	33.4	6.9	241	10	US-09-960-352-7904
13	33	6.9	418	10	US-09-867-701-2763
14	33	6.9	335913	9	US-09-754-853A-2
15	33	6.9	335913	9	US-09-754-853A-3
16	32.6	6.8	683	10	US-09-764-877-914
17	32.6	6.8	2000	9	US-09-938-842A-5219
18	32.4	6.7	2846	9	US-10-008-016-1
19	32.4	6.7	7200	9	US-09-853-450-48

C	20	32.2	6.7	1142	9	US-10-098-841-160	Sequence 160, App
C	21	32.2	6.7	19553	10	US-09-764-847-1425	Sequence 1425, App
C	22	31.8	6.6	411	10	US-09-867-701-7321	Sequence 7321, App
C	23	31.8	6.6	516	10	US-09-960-352-5785	Sequence 5785, App
C	24	31.6	6.6	372	10	US-09-783-590-1292	Sequence 1292, App
C	25	31.6	6.6	1902	10	US-09-349-385-2	Sequence 2, App1
C	26	31.6	6.6	2418	10	US-09-925-300-719	Sequence 719, App
C	27	31.6	6.6	7869	10	US-09-349-385-3	Sequence 3, App1
C	28	31.6	6.6	12886	9	US-10-114-170-14	Sequence 14, App1
C	29	31.4	6.5	4239	10	US-09-815-048-1	Sequence 1, App1
C	30	31.2	6.5	275	10	US-09-867-701-10310	Sequence 10310, A
C	31	31	6.4	2000	9	US-09-938-842A-2799	Sequence 2799, App
C	32	31	6.4	70768	9	US-10-135-322-13	Sequence 13, App1
C	33	30.8	6.4	368	10	US-09-834-975-144	Sequence 144, App
C	34	30.8	6.4	5444	10	US-09-896-617-1	Sequence 1, App1
C	35	30.8	6.4	5444	10	US-09-931-071-1	Sequence 1, App1
C	36	30.8	6.4	32193	9	US-09-764-868-1508	Sequence 1508, App
C	37	30.8	6.4	98865	10	US-09-770-689A-3	Sequence 3, App1
C	38	30.6	6.4	392	10	US-09-833-381-1132	Sequence 1132, App
C	39	30.6	6.4	415	10	US-09-867-701-2926	Sequence 2926, App
C	40	30.6	6.4	520	10	US-09-998-598-434	Sequence 434, App
C	41	30.6	6.4	814	12	US-10-001-870-42	Sequence 42, App1
C	42	30.6	6.4	1205	12	US-10-001-870-43	Sequence 43, App1
C	43	30.6	6.4	2275	10	US-09-764-864-209	Sequence 209, App1
C	44	30.6	6.4	5232	10	US-09-974-592-3	Sequence 3, App1
C	45	30.6	6.4	8726	10	US-09-070-927A-95	Sequence 95, App1

ALIGNMENTS

RESULT 1
US-09-828-456-1
Sequence 1, Application US/09828456
Patent No. US20020051982A1
GENERAL INFORMATION:
APPLICANT: Levy, Stewart et al.
TITLE OF INVENTION: No. US20020051982A1 BLR Molecules Affecting Antibiotic Susc
FILE REFERENCE: PRZ-030
CURRENT APPLICATION NUMBER: US/09/828, 456
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/195,505
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/218,380
PRIOR FILING DATE: 2000-07-04
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 481
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (239)..(361)
US-09-828-456-1
Query Match 100.0%; Score 481; DB 10; Length 481;
Best Local Similarity 100.0%; Pred. No. 5.3e-131;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCCTCATGAGTATGATGCTATTTGACACTATCTTACCCAGCTCAACAGTTT 60
DB 1 TGCCTCATGAGTATGATGCTATTTGACACTATCTTACCCAGCTCAACAGTTT 60
QY 61 AATAACCTGCCAGCAATTAAGGATGTGTTAACTTAAGTCAAAAAATACGATTTTC 120
DB 61 AATAACCTGCCAGCAATTAAGGATGTGTTAACTTAAGTCAAAAAATACGATTTTC 120
QY 121 CAACGACAAAAGCTAATATCGCAAAAACCTCGTAATAATCTTCTGGAGCTATATT 180
DB 121 CAACGACAAAAGCTAATATCGCAAAAACCTCGTAATAATCTTCTGGAGCTATATT 180
QY 181 GCTAAGTAACATTATACCCCTGGAAGTTAATGATCAATCAAGAGATGTGGCTGTAT 240

```
|||||
Db 181 GCTAAGTAACATTTACCCCTGAAGTAAATGATCAATCAAGAGATGTGGCTGAAT 240
QY 241 GAATGCTCTTATTAATTAACAGGTGGATGCTTTCTTGCTTCTTCTTCTTGG 300
Db 241 GAATGCTCTTATTAATTAACAGGTGGATGCTTTCTTGCTTCTTCTTCTTGG 300
QY 301 CGTGGCAGTCACTTTGACAACTATGACCACTGCAAGCAAGTCTTCTTCAACAA 360
Db 301 CGTGGCAGTCACTTTGACAACTATGACCACTGCAAGCAAGTCTTCTTCAACAA 360
QY 361 GTAACTCTGCACTTTGAGACGACATGCTCCGCTCCGGGTGATGTTTCACTTCTG 420
Db 361 GTAACTCTGCACTTTGAGACGACATGCTCCGCTCCGGGTGATGTTTCACTTCTG 420
QY 421 GATATTAACCAAGAAATTTATTTCTTGTCGATGTTGTTGGATTTGTCAGGATAGTA 480
Db 421 GATATTAACCAAGAAATTTATTTCTTGTCGATGTTGTTGGATTTGTCAGGATAGTA 480
QY 481 A 481
Db 481 A 481
```

RESULT 2

```
US-09-938-842A-3476/C
; Sequence 3476, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3476
; LENGTH: 1947
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3476
```

```
Query Match 8.0%; Score 38.6; DB 9; Length 1947;
Best Local Similarity 48.8%; Pred. No. 0.13; 109; Indels 0; Gaps 0;
Matches 104; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
```

```
QY 16 TATGATGCTATTGACACTATCTTACCAACGCTCAACAGTTAATAAACCCTCCACCA 75
Db 16 TATGATGCTATTGACACTATCTTACCAACGCTCAACAGTTAATAAACCCTCCACCA 75
QY 966 TATGATGCTATTGACACTATCTTACCAACGCTCAACAGTTAATAAACCCTCCACCA 907
Db 966 TATGATGCTATTGACACTATCTTACCAACGCTCAACAGTTAATAAACCCTCCACCA 907
QY 76 ATAAGGATGCTTGTACTTACTTAAGCAAAAATAATGCAATTTTCCAGCAAGAAAGCT 135
Db 76 ATAAGGATGCTTGTACTTACTTACTTAAGCAAAAATAATGCAATTTTCCAGCAAGAAAGCT 135
QY 906 TTCAAGTAATTTTGGATTCACATCACCAAAAATAATTAACGAACAGATTAACCAAGT 847
Db 906 TTCAAGTAATTTTGGATTCACATCACCAAAAATAATTAACGAACAGATTAACCAAGT 847
QY 136 AATATCGCAAAAACCTGAGTAAATATCTTGCTGAGCTATATATGCTAAGTAACATTTA 195
Db 136 AATATCGCAAAAACCTGAGTAAATATCTTGCTGAGCTATATATGCTAAGTAACATTTA 195
QY 846 TAAATATCACTTATGTTTTTTTCAAAATATATGATTTGAGATATATGATCGTAATTTGTA 787
Db 846 TAAATATCACTTATGTTTTTTTCAAAATATATGATTTGAGATATATGATCGTAATTTGTA 787
QY 136 CCCCTGAAGTTAATGATCAATCAAGAGAT 228
Db 136 CCCCTGAAGTTAATGATCAATCAAGAGAT 228
QY 786 ACCCATTTTCATCATGAAACAAACAATTAAGTT 754
Db 786 ACCCATTTTCATCATGAAACAAACAATTAAGTT 754
```

RESULT 3

```
US-09-983-965-5260/C
; Sequence 5260, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengding
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(1029)C
; CURRENT APPLICATION NUMBER: US/09/983, 965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465, 231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113, 678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 5260
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 38-LIB34-041-Q1-E1-B6
US-09-983-965-5260
```

```
Query Match 7.7%; Score 36.8; DB 10; Length 332;
Best Local Similarity 51.9%; Pred. No. 0.2;
Matches 83; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

```
QY 103 AAAAAATAGCGAATTTTCCACGACAAAGCTAAATATGCGCAAAAACCTGATTAAT 162
Db 212 AAAAAATAGCGAATTTTCCACGACAAAGCTAAATATGCGCAAAAACCTGATTAAT 162
QY 163 CTGCTGAGCTATTTATGCTAATGATTAATTTACCCCTGATTAATGATCAATCAAG 222
Db 152 TTTCACACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 93
QY 223 AGAGATGAGGCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 262
Db 92 ACATTAGAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53
```

RESULT 4

```
US-09-764-864-792/C
; Sequence 792, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764, 864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 792
; LENGTH: 7053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (154)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-792
```

```
Query Match 7.7%; Score 36.8; DB 10; Length 7053;
Best Local Similarity 51.2%; Pred. No. 0.79;
Matches 86; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
```

```
QY 90 TTAACCTTAAGTCAAAAAATAGCAATTTTCCAAAGCAACAAAGCTAAATATGCAAAA 149
Db 3528 TTAACCTTTGGAATAATATACCTAAGTTTCTGCTTATTTGAGAAATATCTCTTAA 3469
```

```
QY 150 CCTCAGTAAATCTCTGAGCTATTATGCTAAGTAACTTACCCCTGAGTTAA 209
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3468 ACAAAACAAACAAACAGAGAGAGATCATTAACCACTGATATCTTGTATATA 3409
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 210 TGGATCAATCAAGAGAGATGTGGCTGTAATGAATCCTCTTAATGAAT 257
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3408 TAGCTAGTGTAAAGAAATATGATTTGAGTGGTGCATGCAAGTAAGT 3361
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 5

```
US-09-876-889-272
; Sequence 272, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121,466C3
; CURRENT APPLICATION NUMBER: US/09/876,889
; CURRENT FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 272
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(443)
; OTHER INFORMATION: n = A,T,C or G
US-09-876-889-272
```

```
Query Match 7.4%; Score 35.6; DB 10; Length 443;
Best Local Similarity 53.6%; Pred. No. 0.5;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 100 TCAAAAAATGCGAATTTTCCACGACAAAGCTAAATATGCGAAAAACCTCAGTAA 159
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 207 TCAAAACAAATACAAATGCTCTGAGAAAGAAATTTTAAGAACTTAATATAGTAAA 266
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 160 AATCTCTGAGCTATTATGCTAATACATTTACCCCTGAGTAAATGATCATC 219
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 267 AATGCTTCTCTGAATAATTTCTTGAAATTTTAAATGTCAAAATGATGACTCATGC 326
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 220 AAGAGAGATGTGGCTCT 237
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 327 TAAATACATTTGAGGGCT 344
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 6

```
US-09-880-107-3768
; Sequence 3768, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3768
```

```
; LENGTH: 66109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X66401
; NAME/KEY: unsure
; LOCATION: (1)..(66109)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-3768
```

```
Query Match 7.3%; Score 35; DB 10; Length 66109;
Best Local Similarity 47.9%; Pred. No. 7.3;
Matches 101; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
```

```
QY 13 AAGTATGATGGCTATTGACACTATCTTTACCACGCTCAACAGTTTAATACCTGCCA 72
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13691 AAGATATTTCTTAATCTCGCCTCTCTTAATCCCAATCCTCAATGAACTCTTTA 13750
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 73 GCATTAAGGATGTGTTTAACTTAAGTCAAAAAAATAGGAAATTTCCACGACAAAA 132
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13751 ATTAATATTTCTGTTATTTTATTTGAATGAAAAAATTAATATCAACTTTGTCA 13810
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 133 GCTAAATATGCAAAAAACCTCAGTAAATAATCTTGCTGAGCTATTATGCTAAGTAACAT 192
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13811 TGGAAATATAGGAATTAATCTCAGAGATATGATGTTATTAAGTCATTTGACATATTG 13870
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 193 TTACCCCTGAAGTTAATGATCAATCAGA 223
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13871 TTAACTCAGATGCAAAATGCTGAAATTAATA 13901
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 7

```
US-09-880-107-3814
; Sequence 3814, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3814
; LENGTH: 198285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X87344
US-09-880-107-3814
```

```
Query Match 7.3%; Score 35; DB 10; Length 198285;
Best Local Similarity 47.9%; Pred. No. 12;
Matches 101; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
```

```
QY 13 AAGTATGATGGCTATTGACACTATCTTTACCACGCTCAACAGTTTAATACCTGCCA 72
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 89480 AAGATATTTCTTAATCTCGCCTCTCTTAATCCCAATCCTCAATGAACTCTTTA 89539
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 73 GCATTAAGGATGTGTTTAACTTAAGTCAAAAAAATAGGAATTTTCCACGACAAAA 132
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 89540 ATTAATATTTCTGTTATTTTATTTGAATGAAAAAATGAATTAATCAATTTGTCA 89599
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 133 GCTAAATATGCAAAAAACCTCAGTAAATAATCTTGCTGAGCTATTATGCTAAGTAACAT 192
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 89600 TGGAAATATAGGAATTAATCTCAGAGATATGATGTTATTTAAGTCATTTGACATATTG 89659
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Qy	193	TTACCCCTGAAGTTAATGGATCAATCAGA	223
Db	89660	TTAACCCTCAGATGCAAAATGGTGAATTATA	89690

RESULT 8

```

US-09-938-842A-3316/C
: Sequence 3316, Application US/09938842A
: Patent No. US20020160378A1
:
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kreps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCRIPT300-3
:
: CURRENT APPLICATION NUMBER: US/09/938, 842A
: CURRENT FILING DATE: 2001-08-24
:
: PRIOR APPLICATION NUMBER: US 60/227, 866
: PRIOR FILING DATE: 2000-08-24
:
: PRIOR APPLICATION NUMBER: US 60/264, 647
: PRIOR FILING DATE: 2001-01-16
:
: PRIOR APPLICATION NUMBER: US 60/300, 111
: PRIOR FILING DATE: 2001-06-22
:
: NUMBER OF SEQ ID NOS: 5379
:
: SEQ ID NO 3316
:
: LENGTH: 1312
:
: TYPE: DNA
:
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-3316

```

Query Match	7.1%	Score 34.2	DB 9	Length 1512
Best Local Similarity	51.7%	Pred. No. 2.2		
Matches 78; Conservative	0	Mismatches 73	Indels 0	Gaps 0

Qy	93	ACTTACGCAAAAAAATGCGAATTTTCCACGAGCAAAAAAGCTAAATGTGCAAAAAACCT	153
Db	510	AATTTATTTCCAAAGTTATCTAAGAAATTTTCCACTAAGAAAACCTTATCTATGTGACATTC	451b
Qy	153	CAGTAAATCTGTGCGAGCTATTATTGCTAAGTAACATTACCCTGAGTTAATGG	212
Db	450	TTGAACCATCTTAATAACCTAGTTGTCTATTGTAGTTTCGGCAAGCAATGGCAA	391b
Qy	213	ATCATCAAGAGAGTGTGGCTGTAATGAA	243
Db	390	AACAACTTAGTGAGATTGTGTTCTTAACAA	360

RESULT 9

US-09-887-576-859
Sequence 859, Application US/09887576
Patent No. US20020144047A1
GENERAL INFORMATION:
APPLICANT: Budworth, P.
APPLICANT: Brown, D.
APPLICANT: Chang, H.
APPLICANT: Zhu, T.
APPLICANT: Han, B.
APPLICANT: Wang, X.
APPLICANT: Cooper, Bret
TITLE OF INVENTION: Promoters for regulation of plant expression
FILE REFERENCE: 1360.001u51
CURRENT APPLICATION NUMBER: US/09/887, 576
CURRENT FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: US 60/213,848
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/214,087
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/258,692
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 875

```

; SOFTWARE: FastSeq for Windows Version 4.
; SEQ ID NO 859
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
; OS-09-887-576-859

```

Query Match	7.1%	Score 34.2	DB 10	Length 2000
Best Local Similarity	56.8%	Prog. No. 2.6		
Matches 63	Conservative 0	Mismatches 48	Indels 0	Gaps 0

[illegible]

RESULT 10

```

US-10-098-841-110
? Sequence 110. Application US/10098841
? Publication NO. US20020197679A1
? GENERAL INFORMATION:
? APPLICANT: Tang, Y. Tom
? APPLICANT: Liu, Chenghua
? APPLICANT: Asundi, Vinod
? APPLICANT: Xu, Chongjun
? APPLICANT: Zhou, Ping
? APPLICANT: Ma, Yungling
? APPLICANT: Wang, Jian-Rui
? APPLICANT: Zhao, Qing A.
? APPLICANT: Ren, Feiyan
? APPLICANT: Chen, Rui-hong
? APPLICANT: Wang, Dunrui
? APPLICANT: Wang, Zhiwei
? APPLICANT: Weinman, Tom
? APPLICANT: Zhang, Jie
? APPLICANT: Qian, Xiaohong B.
? APPLICANT: Dirmacac, Radoje T.
? TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
? TITLE OF INVENTION: Polypeptides
? FILE REFERENCE: 784C1P2
? CURRENT APPLICATION NUMBER: US/10/098_841
? CURRENT FILING DATE: 2002-03-13
? PRIOR APPLICATION NUMBER: 09/598,042
? PRIOR FILING DATE: 2000-06-20
? PRIOR APPLICATION NUMBER: 09/552,317
? PRIOR FILING DATE: 2000-04-25
? PRIOR APPLICATION NUMBER: 09/488,725
? PRIOR FILING DATE: 2000-01-21
? NUMBER OF SEQ ID NOS: 331
? SOFTWARE: pc_flt_genes Version 1.0
? SEQ ID NO 110
? LENGTH: 2075
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (147)..(2075)
US-10-098-841-110

```

Query Match	7.18;	Score 34.2;	DB 9;	Length 2075;
Best Local Similarity	50.98;	Pred. No. 2.6;		
Matches 81; Conservative	0;	Mismatches 78;	Indels 0;	Gaps 0;

QY 56 ACATTAAATACCTGCCAGCAATAGGAGTGTGGTTTAACCTTAAGTCAAAAAAATAAGGAA 115
| | | | | | | | | | | | | | | |
DB 499 ATTTTAGAANTGCTCAGATATATACAACTTTGAGGGCATACAGGAATATCAAGAAGAA 558
| | | | | | | | | | | | | | | |
QY 116 TTTTCCAAAGAACAAAGCTAATATTCGCAAAAACCCTAGTAAAAAATTTTGCTCGAGCTA 175
| | | | | | | | | | | | | | | |

Db 559 ATATCAATCAAAAGATGATCATGCTATGAAAAAGCTGACTCATACATCTTCATCTTCTC 618
Qy 176 TTATTGCTAGTACATTTTACCCTCGAAGTTAATGAT 214
Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
Db 619 TTATTGCAATACATTAACCGTATGATGTAAGAAAT 657

RESULT 11
US-09-925-302-177/c

; Sequence 177, Application US/09925302
; Patent No. US2002004941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 177
; LENGTH: 4083
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (48)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (157)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-177

Query Match
Best Local Similarity 7.0%; Score 33.6; DB 10; Length 4083;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 53 AACGTTTAATACCTGCGACCATTAAGGATGTTGTTTAACCTTAAGCAAAAATATGCG 112
Db 3725 AAAAGAGTCTTATTCCTCCAAATAGCAAAATTTGTAATTCCTATAGCAAAATATA 3666
Qy 113 GAATTTCCACGAACAAAGCTAAATTCGCAAAA 148
Db 3665 CAATTTTGAATTAATAAAGTTGAGAAATCTTGAAA 3630

RESULT 12

US-09-960-352-7904/c
; Sequence 7904, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7904
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 34-LIB3057-019-Q1-K1-A10
US-09-960-352-7904

Query Match
Best Local Similarity 6.9%; Score 33.4; DB 10; Length 241;
Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 55 CAGTTTAATAACCTGCCAGCAATTAAGGATGTGTTTAACCTTAAGTCAAAAAATAGCGA 114
Db 224 CATATTAATTATACCTAATAGAAATATATGATTTTTTTTAAAAAATTTAAAAAATTAATA 165
Qy 115 ATTTTCCACGACAAAAGCTAAATATCGCAAAAACCTCAGTAAAA 161
Db 164 AAATTAATATATAATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 118

RESULT 13

US-09-867-701-2763
; Sequence 2763, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2763
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-2763

Query Match
Best Local Similarity 6.9%; Score 33; DB 10; Length 418;
Matches 57; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 81 GGATGTGTTTAACCTTAAGTCAAAAAATAGCGAATTTTCCACGACAAAGCTAATA 140
Db 216 GCATTAATTTTAAGCTTTTGAATTAATTAATTAATTCCTCCGCCCAAAAAAAGTA 275
Qy 141 TCGCAAAAACCTCAGTAAAAATCTTGCTGCGACTATT 177
Db 276 TTGTAATACCTTAAGTAAATTAATTTCTAGCTATT 312

RESULT 14

US-09-754-853A-2/c
; Sequence 2, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 2
; LENGTH: 335913
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (45163)..(45314)..(45450)..(45509)..(46941)..(48763)..(48975)..(49573)
; OTHER INFORMATION: Clone ID: 240017_region_03

US-09-754-853A-2

Query Match 6.9%; Score 33; DB 9; Length 335913;
Best Local Similarity 50.3%; Pred. No. 57;
Matches 81; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 102 AAAAAATAGCGAATTTCCACAGACAAAGCTAAATATCGCAAAAACCTCAGTAATAA 161
||| ||| | ||||| ||| | ||||| ||| | ||| ||| | |||
Db 311308 AATATAAATAATATTACACACAACTAAAGCAAAAATACTAAAGACTAAAGAAATA 311249

QY 162 TCTTGCTGAGCTATTAATGCTAAGTAACATTTACCCCTGAAGTTAATGATCAATCAA 221
| | | | | |||| | | | | | | | | | | | | | | |
Db 311248 TTTAAATTAATAGTTATATACCAACAATATAAACAATACAGTTTAAATATCATCTGA 311189

QY 222 GAGAGATGTGGCTGTATGATGATCGTCTTATGATTAACA 262
| | | | | |||| | | | | | | | | | | | | | | |
Db 311188 TGAATAATTAGTATCTAATTAATTATTAACCTTTATATAA 311148

RESULT 15

US-09-754-853A-3/C
; Sequence 3, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 3
; LENGTH: 335913
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46798)..(48763)..(48975)..(49573)
; OTHER INFORMATION: Clone ID: 240017_region_G3
US-09-754-853A-3

Query Match 6.9%; Score 33; DB 9; Length 335913;

Best Local Similarity 50.3%; Pred. No. 57;
Matches 81; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 102 AAAAAATAGCGAATTTCCACAGACAAAGCTAAATATCGCAAAAACCTCAGTAATAA 161
||| ||| | ||||| ||| | ||||| ||| | ||| ||| | |||
Db 311308 AATATAAATAATATTACACACAACTAAAGCAAAAATACTAAAGACTAAAGAAATA 311249

QY 162 TCTTGCTGAGCTATTAATGCTAAGTAACATTTACCCCTGAAGTTAATGATCAATCAA 221
| | | | | |||| | | | | | | | | | | | | | | |
Db 311248 TTTAAATTAATAGTTATATACCAACAATATAAACAATACAGTTTAAATATCATCTGA 311189

QY 222 GAGAGATGTGGCTGTATGATGATCGTCTTATGATTAACA 262
| | | | | |||| | | | | | | | | | | | | | | |
Db 311188 TGAATAATTAGTATCTAATTAATTATTAACCTTTATATAA 311148

Search completed: February 7, 2003, 21:19:15
Job time : 334 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 18:32:14 ; Search time 2441 Seconds

(without alignments)
3191.327 Million cell updates/sec

Title: US-09-828-456-1

Perfect score: 481
Sequence: 1 tgcctctcttgaagtatga.....gtgatgtcagagatgtaa 481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estlom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_juv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	45	9.4	672	AA572059	AA572059 vk95d03.r
2	41.8	8.7	615	AV846478	AV846478 AV846478
3	41.8	8.7	712	AV861290	AV861290 AV861290
4	41.6	8.6	965	AL105557	AL105557 Drosophila
c 5	41.2	8.6	407	BJ364861	BJ364861 BJ364861
6	40.2	8.4	501	AV867201	AV867201 AV867201

7	39	8.1	940	17	BH146288	BH146288
c 8	38.2	7.9	663	17	AZ347477	AZ347477
c 9	38.2	7.9	891	10	BE331179	BE331179
c 10	38	7.9	452	14	H90453	H90453
c 11	38	7.9	633	13	BJ370199	BJ370199
c 12	38	7.9	864	17	BH161244	BH161244
c 13	38	7.9	949	17	AZ547495	AZ547495
c 14	37.8	7.9	485	12	BF083668	BF083668
c 15	37.8	7.9	513	17	BH288998	BH288998
c 16	37.8	7.9	926	17	BH134032	BH134032
c 17	37.6	7.8	630	10	BB631135	BB631135
c 18	37.6	7.8	661	10	AV721985	AV721985
c 19	37.6	7.8	878	17	AZ534410	AZ534410
c 20	37.4	7.8	637	10	AV404587	AV404587
c 21	37.4	7.8	683	17	AZ660149	AZ660149
c 22	37.4	7.8	923	17	CNS0050X	CNS0050X
c 23	37.2	7.7	258	14	BO601966	BO601966
c 24	37.2	7.7	521	17	AO181233	AO181233
c 25	37	7.7	570	13	BJ368191	BJ368191
c 26	37	7.7	619	10	AW180813	AW180813
c 27	37	7.7	797	12	BG434608	BG434608
c 28	37	7.7	1000	17	CNS00C00	CNS00C00
c 29	36.8	7.7	292	9	AI383833	AI383833
c 30	36.8	7.7	295	10	AW515536	AW515536
c 31	36.8	7.7	333	9	AI291645	AI291645
c 32	36.8	7.7	341	9	AI419275	AI419275
c 33	36.8	7.7	372	14	H54495	H54495
c 34	36.8	7.7	382	10	AW576565	AW576565
c 35	36.8	7.7	394	9	AI361450	AI361450
c 36	36.8	7.7	397	9	AA664128	AA664128
c 37	36.8	7.7	421	9	AA983620	AA983620
c 38	36.8	7.7	426	9	AI140942	AI140942
c 39	36.8	7.7	429	9	AI1768981	AI1768981
c 40	36.8	7.7	429	9	AI1818334	AI1818334
c 41	36.8	7.7	430	9	AI985789	AI985789
c 42	36.8	7.7	440	9	AA912460	AA912460
c 43	36.8	7.7	449	9	AA52206.S	AA52206.S
c 44	36.8	7.7	455	9	AI423972	AI423972
c 45	36.8	7.7	465	9	AI802215	AI802215
c 46	36.8	7.7	466	9	AA700428	AA700428

ALIGNMENTS

RESULT 1
AA572059/c
LOCUS
DEFINITION
vk95d03.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA
36.8 KD PROTEIN IN GNTR-GGT INTERGENIC REGION ; , mRNA sequence.

ACCESSION
AA572059
VERSION
AA572059.1 GI:2345254

KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus

REFERENCE
AUTHORS
Marras M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,
1 (Pages 1 to 672)
Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris B.,
Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B.,
Theising B., Wylie T., Lennon G., Soares B., Wilson R. and
Waterston R.

TITLE
JOURNAL
COMMENT
The Washu-HMI Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
Washu-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:551197
High quality sequence stop: 109.
Location/Qualifiers

FEATURES
source

1..672
/organism="Mus musculus"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:962405"
/clone_lib="Knowles Solter mouse blastocyst B1"
/tissue_type="Blastocyst"
/dev_stage="embryo (pre-implantation)"
/lab_host="DH10B"
/note="Organ: embryo; Vector: pSPORT; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally from mRNA prepared
from 800 blastocysts. Primer: SalI(dT):
5'-CGGTGACCGTCGACCGTCTTTTCTTTT-3'. cDNAs were
cloned into the NotI/SalI sites of a pSPORT vector (Life
Technologies). Two different size selections: B1 (larger
inserts) and B3."

BASE COUNT 169 a 160 c 175 g 166 t 2 others
ORIGIN

Query Match 9.4%; Score 45; DB 9; Length 672;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTCTCATTTGAAGTATGATGGCTATTTGACACTATCCTTACC 45
|||||
Db 51 TGCTCTCATTTGAAGTATGATGGCTATTTGACACTATCCTTACC 7

RESULT 2
LOCUS AV846478 615 bp mRNA linear EST 08-NOV-2001
DEFINITION AV846478 Nori Satoh unpublished cDNA library, cleavage stage embryo
Accession AV846478
Version AV846478.1 GI:16825908
Keywords EST.
Source Clona intestinalis.
Organism Clona intestinalis.
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Clona.

REFERENCE 1 (bases 1 to 615)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE Expressed genes in Clona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source Location/Qualifiers

1..615
/organism="Clona intestinalis"
/db_xref="taxon:7719"
/clone="rcic116110"
/clone_lib="Nori Satoh unpublished cDNA library, cleavage
stage embryo"
/tissue_type="whole animal"
/dev_stage="cleavage stage embryo"
/note="Vector: pBluescript SK"

BASE COUNT 220 a 87 c 104 g 203 t 1 others
ORIGIN

Query Match 8.7%; Score 41.8; DB 10; Length 615;
Best Local Similarity 48.2%; Pred. No. 0.99;
Matches 116; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 9 ATTGAAGTATGATGGCTATTTGACACTATCCTTACCACGCTCAACAGTTTAATTAACCT 68

|||||
Db 210 ATAAAGTATGATAAAGGTAATTAACAGATGTCACATCTGACAGTTTAAGTCT 269
QY 69 GCCAGCAATTAAGGATGTTGTTAACTTAAGTCANAAAAATGGAATTTCCAGAAC 128
|||||
Db 270 GGCAGTATTAAGCAAAATTTGCAAAATGATTACACCAATACCCATTATACAAAATGAAA 329
QY 129 AAAAGCTAATATTCGCAAAAACCTCAGTAAATCTTGCTGAGCTATTAATGCTAAGTA 188
|||||
Db 330 AATTTCTAATTTGGTACATGTGTGACATTTATTCGATTAATGATTCAGTTCAAGAAA 389
QY 189 ACATTTACCCCTGAAGTTAATGATCAATCAAGAGATGTGGGCTGTAATGAATCGTC 248
|||||
Db 390 AGTAGAAGTCAAGAAATTAAGTTCAATTAATTTAAGAAAAGTATGATTAATGATCTAC 449
QY 249 TTATT 253
|||||
Db 450 GTCTT 454

RESULT 3
LOCUS AV861290 712 bp mRNA linear EST 08-NOV-2001
DEFINITION AV861290 Nori Satoh unpublished cDNA library, egg Clona
Accession AV861290
Version AV861290.1 GI:16848814
Keywords EST.
Source Clona intestinalis.
Organism Clona intestinalis.
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Clona.

REFERENCE 1 (bases 1 to 712)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE Expressed genes in Clona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source Location/Qualifiers

1..712
/organism="Clona intestinalis"
/db_xref="taxon:7719"
/clone="rcieg34e21"
/clone_lib="Nori Satoh unpublished cDNA library, egg"
/tissue_type="whole animal"
/dev_stage="egg"
/note="Vector: pBluescript SK"

BASE COUNT 237 a 101 c 122 g 242 t 10 others
ORIGIN

Query Match 8.7%; Score 41.8; DB 10; Length 712;
Best Local Similarity 48.2%; Pred. No. 1;
Matches 118; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 9 ATTGAAGTATGATGGCTATTTGACACTATCCTTACCACGCTCAACAGTTTAATTAACCT 68
|||||
Db 178 ATAAAGTATGATAAAGGTAATTAACAGATGTCACATCTGACAGTTTAAGTCT 237
QY 69 GCCAGCAATTAAGGATGTTGTTAACTTAAGTCANAAAAATGGAATTTCCAGAAC 128
|||||
Db 238 GGCAGTATTAAGCAAAATTTGCAAAATGATTACACCAATACCCATTATACAAAATGAAA 297
QY 129 AAAAGCTAATATTCGCAAAAACCTCAGTAAATCTTGCTGAGCTATTAATGCTAAGTA 188
|||||
Db 298 AATTTCTAATTTGGTACATGTGTGACATTTATTCGATTAATGATTCAGTTCAAGAAA 357
QY 189 ACATTTACCCCTGAAGTTAATGATCAATCAAGAGATGTGGGCTGTAATGAATCGTC 248
|||||

FEATURES	location/Qualifiers
SOURCE	1..501
	/organism="Clona intestinalis"
	/db_xref="taxon:7719"
	/clone="rcieg50106"
	/clone_id="Nori Satoh unpublished cDNA library, egg"
	/tissue_type="whole animal"
	/dev_stage="egg"
	/note="Vector: pbluescript SK"
BASE COUNT	177 a 65 c 88 g 163 t 8 others
ORIGIN	
Query Match	8.4%; Score 40.2; DB 10; Length 501;
Best Local Similarity	47.8%; Pred. No. 2.6;
Matches 117; Conservative 0; Mismatches 128; Indels 0; Gaps 0;	
QY	9 ATTGAATGATGAGTCTATTGACACTATCCCTTACCCACGCTCAACGTTTAATACCT 68
DB	150 ATTAAGTAAGTAAAGTAAGGTAATTAACACGATGTTGCACATCTGGCAGTTTAAGTCT 209
QY	69 GCCACCAATTAAGGATGTGTTTACTTAAGCAAAAAAATPCCGAATTTCCACGAAAC 128
DB	210 GGCCTATACACGAAGAAATTTGCCAAATGATTAACCCATATACCAATTTACAAAATGAAA 269
QY	129 AAAAGCTAAATATCCGCAAAAACCTCAGTAAAAATCTTGCAGCTATTATTTGCTAAGTA 188
DB	270 AAATCTTAAATTTGGTACAAATGTGCATTTTATTCGCAATGGAATGTTTCAGTTCAAGAAA 329
QY	169 ACATTACCCCTCGAAGTTAATGATCAATCAATCAAGAGAGATGTGGCTGTAATGATTCGC 248
DB	330 AGTGAGAGCTCAAGAAATTAAGTTCAATTAATTTAAAGAAAGTATGATGATGATCTAC 389
QY	249 TTATT 253
DB	390 GTCTT 394
RESULT 7	
BH146288	940 bp DNA linear GSS 27-AUG-2001
LOCUS	ENTPND9TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION	genomic. DNA sequence.
ACCESSION	BH146288
VERSION	BH146288.1 GI:15301168
KEYWORDS	GSS.
SOURCE	Entamoeba histolytica.
ORGANISM	Entamoeba histolytica
REFERENCE	Eukaryota; Entamoebidae; Entamoeba.
AUTHORS	Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
TITLE	Determination of clone end sequences from Entamoeba histolytica
JOURNAL	HMI:IMSS sheared DNA library (2001)
COMMENT	Unpublished (2001)
	Contact: Brendan J Loftus
	Department of Eukaryotic Genomics
	The Institute for Genomic Research
	9712 Medical Center Dr., Rockville, MD 20850, USA
	Tel: 301 838 0208
	Fax: 301 838 3543
	Email: b.loftus@igr.org
	Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
	DNA library
	Seq primer: M13-Forward
	Class: shotgun
	High quality sequence start: 15
	High quality sequence stop: 703.
FEATURES	location/Qualifiers
SOURCE	1..940
	/organism="Entamoeba histolytica"
	/strain="HMI:IMSS"
	/db_xref="taxon:5759"
	/clone_lib="Entamoeba histolytica Sheared DNA"
	/note="Vector: pHD51; Site: 1; Bst I; Constructed at The

Query Match	8.1%;	Score 39;	DB 17;	Length 940;	
Best Local Similarity	49.7%;	Pred. No. 5.7;			
Matches 99;	Conservative 0;	Mismatches 100;	Indels 0;	Gaps 0;	

BASE COUNT	345 a	89 c	114 g	392 t
ORIGIN				

QY	81	GGATGCTCTTTAACTAAGTCAGAAAATAATAGGAATTTTCCACAGACAGAAAAGCTAATA	140
Db	595	GTAATTTCTTACACAAAAAAGAAACCATGATCTTTTGTTATTAATAATTGGATATAATG	654
OY	141	TCGCAAAAACCTCAGTAATAAATCTTGCTGAGACCTATTATTCCTAAGTAACATTTACCCCC	200
Db	655	TTTCATTTTAATCATATTATTATTATTGACATATTTATTCGCTTTTATTATTTTTAAACCG	714
OY	201	TGAAGTTAAGGATCAATCAAGAGAGATGGCGCTGTAAATGCAATCGCTCTTATGAAATTA	260
Db	715	TTTCATTCATGATGATATAAATAAATAATATGCTAATAATTATATACAAAACAATATGGATTGA	774
OY	261	CAGGTGGATGCTGCTCTGT	279
Db	775	TATTTTCTCTCTTTTGT	793

RESULT	8
AZ347477	
LOCUS	AZ347477 663 bp DNA linear GSS 29-SEP-2000
DEFINITION	I00083005R Mouse 10kb plasmid UUGCLM library Mus musculus genomic clone UUGCIM0083005 R, DNA sequence.
ACCESSION	AZ347477
VERSION	AZ347477.1 GI:10426714
KEYWORDS	GSS.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 663)
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL	unpublished (2000)
COMMENT	Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunne@genetics.utah.edu Insert Length: 10000 Std Error: 0..00 Plate: 0083 row: O column: 05 Seq primer: CACACAGGAACAGCTATGACC Class: plasmid ends High quality sequence stop: 663. Location/Qualifiers 1..663 /organism="Mus musculus" /strain="C57BL/6J"
FEATURES	
Source	

BASE COUNT
ORIGIN

211 a 122 c 121 g 209 t

```

/db_xref="taxon:10090"
/clone="umc1m0083005"
/clone_1lp="mouse 10kb plasmid umc1m library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-
"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g1147321141gb1Ar129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

Query Match	7.9%	Score 38.2	DB 17	Length 663
Best Local Similarity	50.3%	Pred. No. 9.2		
Matches 94	Conservative 0	Mismatches 93	Indels 0	Gaps 0
QY 270	TCGTCCTTGTGCTTTCACATCAATCTTCTTGGCGTGGCAGTGCACATTGACAACTATCAGC	329		
DB 472	TCGGCTGTATGCTGTCACCTTGTAACTCAAGACCTTGTTATCATATGGAATTTGGTGTCA	531		
QY 330	CACCTGAACAGAGTCTCTCGGTACACACAAGTAAGCTCTGCACCTTGTGGAGCGACATGC	389		
DB 532	CAAACTCAACTATGTCCTATGTATGATGCAAAAATTTTTCMAACCACTGTGGATACACATTT	591		
QY 390	TGCCCGTCGGGTGATCTTTTCACATCTGCGGATTTTAACACAGAAATTTTATCTGT	449		
DB 552	TGCCAGTCCAGCTAAATATATTATCTTGGCCATTTTTAAGAAAAAGATTTCTTTTAA	651		
QY 450	TGCGATGT 456			
DB 652	ATTATGT 658			

RESULT 9	BE531179/c	891 bp	mRNA	linear	EST 09-AUG-2000
LOCUS	BE531179	891 bp	mRNA	linear	EST 09-AUG-2000
DEFINITION	601278311F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610347 5', mRNA sequence.				
ACCESSION	BE531179				
VERSION	BE531179.1	GI:9759915			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 (bases 1 to 891)				
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgabs-remail.nih.gov				
	Tissue Procurement: ATCC				
	cDNA Library Preparation: Ling Hong/Rubin Laboratory				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:				
	http://image.lnl.gov				

BASE COUNT	ORIGIN
162 a	277 c
292 g	160 t

Query Match	Best Local Similarity	7.9%; Score 38.2; DB 10; Length 891;
Matches 76; Conservative 0; Mismatches 63; Indels 0; Gaps 0;		
QY 245	CGCTTATTGAATTAACAGGTGGATCCCTTGTCTGTTTCAGTCATTCCTTGGCGTG	304
Db 746	CGCTCATATGAGGGGTCGAGGAAAGTATGTCATTCATTCAGGGGATCCACACTTGGCCTC	687
QY 305	GGCAGTCACATTTGACAACTATACAGCCACCTGACAGAGTGGTTGGGTACACACAGTAA	364
Db 686	CAGCTCCACCTTGGTCATATGAGCAGCATCCACATCCTTCTTCAGCACCAACAACTCAT	627
QY 365	GCTCGCAGCTGTGGAGCG	383
Db 626	TCTCAGCAGCTGTGGCGG	608
RESULT 10		
LOCUS H90453		
DEFINITION	H90453 452 bp mRNA linear EST 28-NOV-1995	
ACCESSION	U01005.F1 Soares fetal liver spleen INFLS Homo sapiens CDNA clone	
VERSION	H90453	
KEYWORDS	H90453.1 GI:1080883	
SOURCE	EST.	
ORGANISM	human.	
REFERENCE	Homo sapiens	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
	1 (bases 1 to 452)	
	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman	
	,M., Hulman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,	
	Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Treaskis,E., Waterston	
	,R., Williamson,A., Wobldmann,P. and Wilson,R.	
TITLE	The Wash-Merck EST Project	
JOURNAL	Unpublished (1995)	
COMMENT	Contact: Wilson RK	
	Washington University School of Medicine	
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108	
	Tel: 314 286 1800	
	Fax: 314 286 1810	
	Email: est@watson.wustl.edu	
	Insert Size: 2214	
	High quality sequence stops: 210	
	Source: IMAGE Consortium, LNL	
	This clone is available royalty-free through LNL; contact the	
	IMAGE Consortium (info@image.llnl.gov) for further information.	
	Insert Length: 2214	
	Seq primer: M13RPI	
	High quality sequence stop: 210.	
FEATURES		
source	1..452	
	Location/Qualifiers	

```

/organism="Homo sapiens"
/db_xref="GDB:3790581"
/lab_xref="taxon:9606"
/clone_image="241448"
/clone_id="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: pTV73D (Pharmacia) with a modified polylinker; Site.1: Pac I; Site.2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGCGACAGATTAAATTAAAGACTCTTTTCTTTTCTTTTCTT 3'], (double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTV73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

```

Query Match	7.9%	Score 38;	DB 14;	Length 452;
Best Local Similarity	51.2%	Pred. No. 10;		
Matches	86;	Conservative	0;	Mismatches 82; Indels 0; Gaps 0;
QY	90	TTAACTTAAGTC	CAAAAATATGCGAATTTTCC	CAAGCAAAAAGCTAATATCGCAAAA 149
Db	41	TTAACTTTTGCAAAATATATACCTAAGGTTTTCTGCGCTATTATGAGCAAAATTTCC	TACAA 100	
QY	150	CCTCAGTAAAAATCTTGCTGGAGCTATTATTGCTTAAGTACATTTAAACCCCTGAAAGTTAA 209		
Db	101	ACAAACAACAACAACAAGAGAGACATCTATTAAACACTGTATATCTTTGGTATATPAA 160		
QY	210	TGAGTCATCAAGAGAGATGCGGCTGTATGCAATCGGCTATTGTAAT 257		
Db	161	TAGGTACAGTGTAAAGAAATATGATTTTGGAGTGTCATGCAAGTAAGTACT 208		

RESULT 11
 BJ370199
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

BJ370199 633 bp mRNA linear EST 08-MAR-2002
 BJ370199 Dictyostellium discoideum cDNA library, CF Dictyostellium
 discoideum cDNA clone ddc5c12 5', mRNA sequence.
 BJ370199
 BJ370199.1 GI:19279582
 EST.
 Dictyostellium discoideum.
 Dictyostellium discoideum
 Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
 1 (bases 1 to 633)
 Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
 Full length cDNA of Dictyostellium at the culmination
 stage
 unpublished (2002)
 Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin1@genes.nig.ac.jp.

FEATURES	source	Location/Qualifiers
		1..633
		/organism="Dictyostelium discoideum"
		/strain="AX4"
		/db_xref="taxon:44689"
		/clone="ddc53c12"
		/clone_11b="Dictyostelium discoideum cDNA library, CF"
		/sex="mat A"
		/dev_stage="Culmination stage"
BASE COUNT	226 a	73 c 116 g 217 t 1 others
ORIGIN		
Query Match		7.9%: Score 38; DB 13; Length 633;

	Best Local Similarity	47.8%	Pred. No.	10;
Matches	110;	Conservative	0;	Matches 120;
			Indels	Gaps
			0;	0;
QY	GCATTAAGCGATTGTTACTTAACTTAAGTCACAAAAAATAAGCATTTTCCACAGACAANA	132		
Db	GCACCAAGAGAAGATCTTGATGTATTAAATGTATGTGGGAATGGATTAATATACCATA	165		
QY	133 GCTAAATATTCGCAAAAACCTCAGTAAATAATCTTGCTGGAGCATTTATATGCTAATTAACAT	192		
Db	166 GGATGTAATTCACACTAATTAATAGTAGTAATTAATGAGCAGAGATGGATGATGATCAT	225		
QY	193 TTACCCCCCTGAAGTTAATGGATCAATCAAGAGAGATGTGGCGTGTAAATGATCGTTAT	252		
Db	226 TATACAAACAATGTGATGGTATGTTAATTTAAATTTTCAAAATAAATGAATTAATGGAT	285		
QY	253 TGAATTAACAGTTGGATCGTCTCTGTGGTTCACTCATTTCTCTGGCG	302		
Db	286 TTAATTAATTAATTTGGCAGAGATCAAGTGACATATTTTACGTGGTTGGCG	335		

RESULT	12
LOCUS	BH161244
DEFINITION	BH161244 864 bp DNA linear GSS 24-SEP-2001
ACCESSION	ENTRXSITF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.
VERSION	BH161244
KEYWORDS	BH161244..1 GI:15734682
SOURCE	GSS.
ORGANISM	Entamoeba histolytica. Entamoeba histolytica Eukaryota; Eutrombozoidea; Eumetazoa; Alveolata; Excavata; Amoebozoa; Eukarya;
REFERENCE	1 (bases 1 to 864)
AUTHORS	Lofthus,B., Wang,Z., Van Aken,S. and Fraser,C.
TITLE	Determination of clone end sequences from Entamoeba histolytica HMI:IMSS sheared DNA library (2001)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Brendan J Lofthus Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0208 Fax: 301 838 3543 Email: bjl@lofustelgr.org Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library Seq primer: M13-Forward Class: shotgun High quality sequence start: 5 High quality sequence stop: 728.

```

FEATURES
    source
        location/Qualifiers
            1..864
                /organism="Entamoeba histolytica"
                /strain="HML:IMSS"
                /db_xref="taxon:5759"
                /clone_lib="Entamoeba histolytica Sheared DNA"
                /note="Vector: pHOSt1; Site_1: Bst I; Constructed at The
                Institute for Genomic Research (TIGR), Rockville, MD.
                Genomic DNA isolated from broth cultures of E. histolytica
                using a method described by Clark and Diamond (Clark,
                C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
                method for isolate identification. Exp. Parasitol.
                77:450.). The DNA was mechanically sheared to give a
                tight size distribution (~2 kb). The v + i method used for
                the library construction is described in detail in Smith
                H.O. and Venter, J.C. (Making small insert libraries for
                whole genome shotgun sequencing projects. In Genome
                Sequencing: A Practical Approach", eds. M. Vaudin and B.
                Barrell, Oxford University Press, 1999)."
```

BASE COUNT	341 a	94 c	99 g	330 t
ORIGIN				
Query Match	7.9%;	Score 38;	DB 17;	Length 864;
Best Local Similarity	48.2%;	Pred No. 11;		

RESULT	13
AZ547495/c	
LOCUS	AZ547495 949 bp DNA linear GSS 14-NOV-2000
DEFINITION	ENTEMM44TtB Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.
ACCESSION	AZ547495
VERSION	AZ547495.1 GI:11170256
KEYWORDS	GSS.
SOURCE	Entamoeba histolytica.
ORGANISM	Entamoeba histolytica
REFERENCE	Eukaryota; Entamoebidae; Entamoeba.
AUTHORS	1 (bases 1 to 949)
TITLE	Lofthus,B., Van Aken,S. and Fraser,C. Determination of clone end sequences from Entamoeba histolytica HMI:IMS shared DNA library
JOURNAL	Unpublished (2000)
COMMENT	Contact: Brendan J Lofthus

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel.: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@igr.org
Clones are derived from the *Entamoeba histolytica* HM1:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 33
High quality sequence stop: 819.
Location/Qualifiers
1..949

```

/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broch cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
Whole genome shotgun sequencing projects. In Genome
Sequencing: A practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

```

Query Match	7.98;	Score 38;	DB 17;	Length 949;
Best Local Similarity	48.28;	Pred. No. 11;		
Matches 107;	Conservative	0;	Mismatches 115;	Indels 0;
				Gaps 0;

RESULT	14
BF083668	
LOCUS	BF083668 485 bp mRNA linear EST 18-OCT-2000
DEFINITION	MR-C70456-130900-001-h12 C70456 Homo sapiens cDNA, mRNA sequence.
ACCESSION	BF083668
VERSION	BF083668.1 GI:10877407
KEYWORDS	EST.
SOURCE	human.

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Homo sapiens	1 (bases 1 to 485)	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Britones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordini, S., Costa, F. F., Goldmann, G. H., Carvalho, A. F., Matsumura, A., Bala, G. S., Simpson, D. H., Brunstein, A., deOliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpc@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-MR2-CR0456-130>)
 900-001-1126r3-2000-09-13c4t4-1)
 Seq primer: pnc 18 forward
 High quality sequence start: 28
 High quality sequence stop: 358.
 Location/Qualifiers
 1..485

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CR0456"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low

```

BASE COUNT	158	a	101	c	121	g	105	t
ORIGIN	stringency conditions.							
'Query Match	7.98; Score 37.8; DB 12; Length 485							
Best Local Similarity	55.88; Pred. No. 12;							

	Matches	72:	Conservative	0:	Mismatches	57:	Indels	0:	Gaps	0:
OY	46	CACGCTCAACAGTTTAATTAACCTGCCAGCATAAGGATGTGGTTAACTTAAGTCAAAA	105							
Db	226	CAAGGTGCACAGGATATTAACCTGCACCGAGACACTCAGGTGTGGCCAAAAAAAAAAAAA	285							
OY	106	AAATATGCAATTCTTCCAACGAACAACAAAGCTTAATATGCCAAAACCTCGTAAAAATCTT	165							
Db	286	AAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTCTTTTAAATGTC	345							
OY	166	GCTGGAGCT	174							
Db	346	TTGGGGGCT	354							

RESULT	15	
LOCUS	BH288998	
DEFINITION	BH288998	513 bp DNA linear GSS 30-NOV-2001
ACCESSION	CH230-2L9, TV CHORI-230 Segment 1	Rattus norvegicus genomic clone
VERSION	BH288998	
KEYWORDS	BH288998.1	GI:17201406
SOURCE	GSS.	
ORGANISM	Norway rat.	
	Rattus norvegicus	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 513)	Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de Jong, P. and Fraser, C. M.	Rat PAC End Sequences from Library CHOR-230 Ecort segment	Unpublished (1999)	Other_GSSs: CH230-219..TJ

Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat330.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@tigr.org)
Clones may be purchased from BACAC Resources
(http://www.chori.org/bacpac/orientation_information.htm). BAC end
page: http://www.tigr.org/tadb/bac_ends/rat/bac_end_intro.html
Plate: 2 row: L column: 9
Seq primer: T7
Class: BAC ends.

FEATURES	Location/Qualifiers
source	1. .513

```

/organism="Rattus norvegicus"
/strain="BN/SAHsd/MCw"
/db_xref="taxon:10116"
/clone="CH230-2L9"
/clone_id="CHOR1-230 Segment 1"
/sex="Female"
/cell_type="Brain"
/notes="Vector: PTARPA2.1; Site_1: ECORI; Site_2: ECORI
CHOR1-230 Rat (BN/SAHsd/MCw) BAC library produced by
Plater de Jong"

```

Query Match	7.98;	Score 37.8;	DB 17;	Length 513;
Best Local Similarity	48.48;	Pred. No. 12;		
Matches 105;	Conservative	0;	Mismatches 112;	Indels 0;
				Gaps 0;

Qy 5 TCATGAGTATGATGGCTATTGTGACACTATCCCTTTACCACCGCTCAACAGTTTAATA 64
|| ||| || | || | || | || | || | || |
Db 78 TCAATTCATTTTTTTATTCATTAACAATACATTTTTATAACACAAAAGATTAAATGACT 137

Qy	65	ACGTCGCGAGCAATTAAGGATCTCTGTTTAACCTAAGTCGCAAAAAAATACGATTTCCAA	124
Db	138	AAAAATATTATTAATAATAGTTGTATGTGAAAAATGTAACAAGACATATTAATATGCAAC	197
Qy	125	GACAAAAGCTAAATATATGCAAAAACCGCAGTAAATAATCTGCGAGCATATTTGCTA	184
Db	198	AACAAAAGAAATTTCAATTAAATATCTAACAAAATAATTTCTTAACATCATATCTTC	257
Qy	185	AGTAACATTACCCCTGAAGTAAATGATGATCAATCAA	221
Db	258	ACGACATTTTAGAAGCACAGATGATTTCTCTTTCTA	294

Search completed: February 7, 2003, 21:11:55
Job time : 2449 secs